

OM of: US-09-303-518d-127 to: PIR_71.* out_format : pfs
 Date: Jun 30, 2002 7:40 AM
 About: Results were produced by the GenCore software, version 4.5,
 Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+nt2p.model -DEV=xlh
 -O=/cgn2_1/uspt0_spool/US09303518/runat_28062002_142713_4317/app_query.fasta_1.23501
 -DB=PIR_71 -QFMT=fastan -SUFFIX=rpr -GAPOP=12.000 -GAPEXT=4.000
 -MINWATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000 -GAPOP=4.500
 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -XGAPOP=6.000
 -DELEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
 -LIST=100 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
 -ALIGN=45 -MODE=LOCAL -OUTFMT=pfs -NORW=ext -HEAPSIZ=500
 -MINLEN=0 -MAXLEN=2000000000 -USER=US09303518 -CGNJ_1_1092
 -NCPD=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
 -NO_TMPXY -WAIT -THREADS=1

Search information block:

Query: US-09-303-518d-127

Query length: 1344

Database: PIR_71.*

Database sequences: 283138

Database length: 96089334

Search time (sec): 504.980000

score_list:

Sequence	Strd Orig	zScore	Escore	Len	Documentation
pir2:A81919	+ 2211.00	3391.60	1.8e-181	447	! probable sodium-translocating N
pir2:D81185	+ 2177.00	3339.20	1.5e-178	447	! sodium-translocating NADH dehyd
pir2:T64002	+ 1574.00	2409.91	8.6e-127	447	! sodium-translocating NADH dehyd
pir2:AG0393	+ 1531.00	2343.64	4.2e-123	447	! NADH dehydrogenase (ubiquinone)
pir2:S51015	+ 1422.00	2176.45	8.7e-114	446	! sodium-translocating NADH dehyd
pir2:G82094	+ 1412.50	2160.69	6.3e-113	462	! sodium-translocating NADH dehyd
pir2:H83272	+ 1298.00	1984.61	4.2e-103	445	! sodium-translocating NADH dehyd
pir2:D86583	+ 447.50	673.41	4.4e-30	467	! ubiquinone oxidoreductase, alpha
pir2:T72040	+ 441.00	673.41	4.4e-30	467	! probable sodium-translocating N
pir2:C81751	+ 441.00	663.44	1.6e-29	465	! probable sodium-translocating N
pir2:T71489	+ 425.00	638.78	3.7e-28	465	! sodium-translocating NADH dehyd
pir2:E82252	+ 137.00	220.41	4.3e-05	801	! rnfC-related protein VC1015 [im
pir2:S69703	+ 132.50	177.19	0.0053	673	! HKR1 protein precursor - yeast
pir2:AE0692	+ 131.00	184.37	0.0080	673	! probable NADH reducing dehydrog
pir2:S59310	+ 127.50	179.05	0.0132	528	! gastric mucin (clone PGM-2A) -
pir2:T34434	+ 123.50	164.81	0.0352	232	! hypothetical protein K06A9.1a
pir2:A40670	+ 123.50	162.75	0.0414	1367	! nuclear envelope protein POM 1
pir2:S48478	+ 122.50	162.75	0.0414	1367	! glucan 1,4-alpha-glucosidase C
pir2:T45025	+ 121.00	168.92	0.0587	3570	! mucin MUC5B, tracheobronchial
pir2:D39903	+ 121.00	168.92	0.0587	3570	! serine-rich protein - fission y
pir2:D83208	+ 121.00	165.27	0.0529	774	! probable ferredoxin PA3491 [imp
pir2:E72398	+ 120.00	169.04	0.0559	451	! hypothetical protein TM0244 - T
pir2:T34433	+ 118.00	157.81	0.1032	1032	! hypothetical protein K06A9.1a
pir2:G84348	+ 117.00	163.52	0.1037	494	! hypothetical protein Vng1983h
pir2:T39903	+ 116.50	158.05	0.1297	797	! glycoprotein X precursor - equi
pir2:B64136	+ 116.50	157.78	0.1307	819	! rnfC protein homolog - Haemophi
pir2:T02345	+ 116.50	150.08	0.1604	1791	! hypothetical protein KIAA0324
pir2:T05352	+ 110.00	147.31	0.4778	857	! hypothetical protein F8B4.120 -
pir2:S63399	+ 109.50	152.73	0.4473	457	! probable membrane protein YNR0
pir2:T64184	+ 109.50	151.48	0.4625	519	! UDP-N-acetylmuramoylalanyl-D-gl
pir2:S39893	+ 109.50	151.48	0.4625	519	! rnfC protein - Rhodobacter caps
pir2:S55316	+ 109.00	155.56	0.4486	317	! mucin (clone PGM-2B) - pig
pir2:T52257	+ 109.00	148.79	0.5374	631	! epistatin - mouse
pir2:C39135	+ 108.00	150.86	0.5947	437	! hypothetical protein 3 (gyrB re
pir2:T45134	+ 108.00	148.98	0.6252	529	! hypothetical protein [imported]
pir2:T45462	+ 108.00	144.13	0.7115	866	! membrane glycoprotein [imported]
pir2:T14578	+ 108.00	140.76	0.7782	1219	! nucleoporin Nup153 homolog - A
pir2:S52137	+ 107.50	151.57	0.6311	376	! MID2 protein - yeast (Saccharom
pir2:A41258	+ 107.50	145.11	0.7496	725	! a-agglutinin core protein AGA1

seq_name: pir2:A81919	107.00	147.90	0.7526	505
pir2:B46629	+ 106.50	150.68	0.7558	352
pir2:C83451	+ 106.50	147.60	0.8203	481
pir1:PWFrBL	+ 106.50	141.81	0.9573	867
pir2:T45463	+ 106.50	138.01	1.06	1275
pir2:T33369	+ 105.50	144.32	1.05	574
pir2:S74849	+ 105.50	134.18	1.37	1609
pir2:S25345	+ 105.00	146.28	1.07	435
pir2:F89898	+ 105.00	142.63	2.63	13288
pir2:T03099	+ 103.50	128.60	2.18	2073
pir2:T43311	+ 103.50	128.60	2.18	2073
pir2:T33207	+ 103.00	143.20	1.60	435
pir2:A69632	+ 103.00	142.21	1.64	481
pir2:S22509	+ 103.00	139.17	1.78	555
pir2:G75582	+ 103.00	137.97	1.83	740
pir2:G64919	+ 103.00	137.97	1.83	740
pir2:F87619	+ 103.00	137.55	1.85	772
pir2:B90921	+ 102.50	138.03	1.98	680
pir2:S21323	+ 102.50	137.83	1.99	694
pir2:JCA301	+ 102.00	143.09	1.87	376
pir2:C75580	+ 102.00	139.62	2.05	535
pir1:I39730	+ 102.00	139.48	2.06	543
pir2:S35047	+ 102.00	138.83	2.10	580
pir2:T43481	+ 102.00	136.30	2.24	750
pir2:T42614	+ 101.50	140.91	2.15	434
pir2:AB1597	+ 101.50	118.17	3.93	4377
pir2:A55575	+ 101.00	138.46	3.07	1161
pir2:S57180	+ 100.50	138.52	2.67	473
pir2:A84943	+ 100.50	137.90	2.72	504
pir2:AD1994	+ 100.50	129.94	3.36	1131
pir2:T41144	+ 100.50	124.16	3.92	2035
pir2:AA0718	+ 100.00	139.26	2.84	406
pir2:T38928	+ 100.00	138.58	2.89	435
pir2:B83958	+ 100.00	137.07	3.01	507
pir2:AB9333	+ 100.00	137.07	3.01	507
pir2:AB9230	+ 100.00	136.38	3.06	544
pir2:AF1900	+ 100.00	131.72	3.47	873
pir2:F96615	+ 99.50	135.48	3.39	551
pir2:S64314	+ 99.50	132.96	4.61	1777
pir2:T34369	+ 99.00	135.13	4.83	1459
pir2:AI2488	+ 99.00	121.14	5.37	2187
pir2:T30826	+ 98.50	136.29	3.88	434
pir2:AD1234	+ 98.50	130.32	4.55	796
pir2:T21460	+ 98.00	135.48	5.60	1203
pir2:S26650	+ 98.00	132.67	5.72	1306
pir2:S23378	+ 97.50	132.17	5.07	564
pir2:AH2328	+ 97.50	131.48	5.16	605
pir2:S48940	+ 97.50	129.10	5.50	770
pir2:F96660	+ 97.00	134.73	5.12	402
pir2:G70718	+ 97.00	131.95	5.51	533
pir2:E86185	+ 97.00	130.23	5.77	635
pir1:S75536	+ 97.00	128.53	6.04	755
pir2:F54777	+ 97.00	125.73	6.50	1003
pir2:S23441	+ 97.00	123.66	6.87	1238
pir1:ACUHPR	+ 97.00	119.18	7.74	1952
pir2:T03465	+ 96.50	141.23	4.65	192
pir2:T48814	+ 96.50	136.33	5.30	316
pir2:F71169	+ 96.50	134.31	5.60	388
pir2:G48187	+ 96.50	127.36	6.73	786
pir2:T10963	-			
pir2:T16509	-			

seq_name: pir2:A81919

seq_documentation_block:

probable sodium-translocating NADH dehydrogenase (ubiquinone) (EC 1.6.5.-) chain A NM
 C:Species: Neisseria meningitidis
 C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 16-Feb-2001
 C:Accession: A81919
 R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo
 Nature 404, 502-506, 2000
 A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491
 A:Reference number: A81775; MUID:20222556

A:Accession: A81919
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-447 <PAR>
A:Cross-references: GB:AL162754; GB:AL157959; NID:g7379424; PID:CA884035.1; PID:g737947
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: ugrA; NNA0752
C:Superfamily: Vibrio alginolyticus sodium-translocating NADH dehydrogenase (ubiquinone)
C:Keywords: NAD; oxidoreductase

alignment_scores:
Quality: 2211.00 Length: 447
Ratio: 5.059 Gaps: 0
Percent Similarity: 97.763 Percent Identity: 97.092

alignment_block:

US-09-303-518D-127 x A81919 ..

Align seg 1/1 to: A81919 from: 1 to: 447

```
1 ATGATTAATAAATAAGGCTCTAAAGCTCCCATCGCGGCGACACGGA 50
1 MetileLysileLysLysGlyLeuAsnLeuProileAlaGlyArgProG1 17
51 GCAAGTCATTTATGACGGGCGGCTCATACCGAGTCGCGTTCGTTGGCG 100
17 uGlnValileTyrAspGlyProValIleThrGluValAlaLeuLeuGlyG 34
101 AAGAATATGCGGCTATGCGGCGCTTNGATGAAAGTCAAGGAGGCGATGCC 150
34 LuGluTyrAlaGlyMetArgProSerMetLysValLysGluGlyAspAla 50
151 GTCAAAAAGGCCAAGTCTGCTTTTGAAGACAAAGNATCCGGGCGTGGT 200
51 ValLysLysGlyGlnValLeuPheGluAspLysLysAsnProGlyValVa 67
201 GTTTACCGCCGCTTTTCAGGCAAAATCGCGCATCCATCGCGCGGCAAA 250
67 lPheThrAlaProValSerGlyLysIleAlaAlaIleHisArgGlyGluL 84
251 AGCGGCTACTTCAGTCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
84 ysArgValLeuGlnSerValValIleAlaValGluGlyAsnAspGluIle 100
301 GAGTTCGAACGCTACGCGCGCGAGCGGTGGCAAACTTAAGCGCGGANGA 350
101 GluPheGluArgTyrAlaProGluAlaLeuAlaAsnLeuSerGlyGluG1 117
351 ANTNNNGCAATCTGATCCCAATCCGTTTGTGGACTCGCTGCTGCTANCC 400
117 uValArgArgAsnLeuIleGlnSerGlyLeuTyrThrAlaLeuArgThrA 134
401 GTCCGTTTCAGCAAAATCCCTGCGCTGATGCGGAGCGGCTTCGCGCATCTC 450
134 rgProPheSerLysileProAlaValAspAlaGluProPheAlaIlePhe 150
451 GTCATATGCGATGGACACCAATCCCTGCGGCGGCGGCGGCGGCGGCGGAT 500
151 ValAsnAlaMetAspThrAsnProLeuAlaAlaAspProValValIle 167
501 CAAGAAGCCGCGGANGATTTTCAGACGANGTNTGCTGCTGATGAGCGGTT 550
167 eLysGluAlaAlaGluAspPheArgArgGlyLeuLeuValLeuSerArgL 184
551 TGACCGAGCGTAAATCCATGCTGTGTGAAGGAGCGTGGCGGAGAGCGTCCG 600
184 euThrGluArgLysileHisValCysLysAlaAlaGlyAlaAspValPro 200
601 TCTGNAATGCTGCAACATCGAACACATGAATTCGGCGCGCGCGCATCC 650
201 SerGluAsnAlaAlaAsnIleGluThrHisGluPheGlyGlyProHisPr 217
```

```
651 GGCGGTTTGTAGTGGCAGCGCATTCATTTTCATTCAGCGGTCGGTGCAA 700
217 oblaGlyLeuSerGlyThrHisIleHisPheIleGluProValGlyAlaA 234
701 ACAAAACCGCTTTGGACCATCAATATCAAGATGTAATTCGCCATCGGAGT 750
234 snLysThrValTrpThrIleAsnTyrGlnAspValIleAlaIleGlyArg 250
751 TTGTTTGAACAGCGGCTCTGACACCGCGCTGATTCGCTTCTTGGGTGG 800
251 LeuPheAlaThrGlyArgLeuAsnThrGluArgValIleAlaLeuGlyG1 267
801 TTCTCAAGTCAACAAACCGCGCTCTTTCGCTACCGTTTGGGTGGCGAAG 850
267 ysSerGlnValAsnLysProArgLeuLeuArgThrValLeuGlyAlaLysV 284
851 TATCGCAAAATTAATCGCGGCGGAATTTGGTTGACGACAGACACCGCGTGA 900
284 alSerGlnIleThrAlaGlyGluLeuValAspAlaAsnArgValIle 300
901 TCCGTTTCGTTTCAACGCGCGGATTCACACAGCGCGCGCATTCATTT 950
301 SerGlySerValLeuAsnGlyAlaIleThrGlnGlyAlaHisAspTyrFle 317
951 GGGACGCTACCAATCAAGATTCGCTTATCGAAGAGCGCGCGACCAAG 1000
317 uGlyArgTyrHisAsnGlnIleSerValIleGluGlyArgSerLysG 334
1001 AGCTGTTCGGTGGTGGTGGCGCGGAGCGGACAAATATCCATCAGCGT 1050
334 LuLeuPheGlyTrpValAlaProGlnProAspLysTyrSerIleThrArg 350
1051 AGCACCTCGCGCATTCCTGAAACAAACTCTTCAAGTTTCACGACAGC 1100
351 ThrThrLeuGlyHisPheLeuLysAsnLysLeuPheLysPheThrFAl 367
1101 CGTCAACGTTGGCGCGCATTCGCTGCGGATTCGCTGCTTACGACGCG 1150
367 aValAsnGlyGlyAspArgAlaMetValProIleGlyThrTyrGluArgV 384
1151 TAATGCGCTAGACATCCCTGCGCTACCTGCTTTCGCGGATTTAATCGTC 1200
384 alMetProLeuAspIleLeuProThrLeuLeuLeuArgAspLeuIleVal 400
1201 GCGGATACGACGCGCGCAACATTCGCTTGGTTCCTTGAATTCGACGA 1250
401 GlyAspThrAspSerAlaGlnAlaLeuGlyCysLeuGluLeuAspGluG1 417
1251 AGACCTCGCTTTGTGACGCTTCGCTGCGCGGCGCAAAATACGAATANGCC 1300
417 uAspLeuAlaLeuCysSerPheValCysProGlyLysTyrGluTyrGlyP 434
1301 CGCTGTTCGTTAAGTGTGGAAACCTTGAGAAGGAAGGC 1341
434 roLeuLeuArgLysValLeuGluThrIleGluLysGluGly 447
```

seq_name: pir2:D81185

seq_documentation_block:

sodium-translocating NADH dehydrogenase (ubiquinone) (EC 1.6.5.-) alpha chain NMB0569
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 02-Mar-2001
C:Accession: D81185
R:Tetelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, R.; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.; Ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755
A:Accession: D81185
A:Status: preliminary

A: Molecule type: DNA
 A: Residues: 1-447 <TET>
 A: Cross-references: GB:AE002412; GB:AE002098; NID:97225783; PIDN:AAF40997.1; PID:9722579
 A: Experimental source: serogroup B, strain MC58
 C: Genetics:
 A: Gene: NMB0569
 C: Superfamily: vibrio alginolyticus sodium-translocating NADH dehydrogenase (ubiquinone)
 C: Keywords: oxidoreductase

alignment_scores:
 Quality: 2177.00 Length: 447
 Ratio: 4.982 Gaps: 0
 Percent Similarity: 97.763 Percent Identity: 94.855
 alignment_block:
 US-09-303-518d-127 x D81185 ..
 Align seg 1/1 to: D81185 from: 1 to: 447

```

1 ATGATTAATAACAAAAAGGTCTAAACCTGCCCATCGCGGCAGACCGGA 50
|||||
1 MetileLysIleLysLysLysLysLysLeuAsnLeuProIleAlaGlyArgProG 17
51 GCAAGTCATTATGACGGCCCGCTCATACCGAAGTCGCGTGGCG 100
|||||
17 uGlnAlaValIyrAspGlyProAlaIleThrGluValAlaLeuLeuGlyG 34
101 AAGAAATATGCCGGTATGCCGCCCTNGATGAAAGTCAAGGAAGCGGATGCC 150
|||||
34 LuGluTyrAlaGlyMetArgProSerMetLysValLysGluGlyAspAla 50
151 GTCAAAAGGCCAAGTCTGTTGAGACAAAGAAAGNATCGCGCGTGGT 200
|||||
51 ValLysLysGlyGlnValLeuPheGluAspLysLysAsnProGlyValva 67
201 GTTTACCGCGCCNGTTTCAGGCAAAATCGCGCCCATCCATCGCGCGCAA 250
|||||
67 lPheThrAlaProAlaSerGlyLysIleAlaAlaIleHisArgGlyGlu 84
251 AGCCGCTACTTCAGTCGGTGTGATTCGGCTTGAAGCAACGACGAAATC 300
|||||
84 ysArgValLeuGlnSerValValIleAlaValGluGlyAsnAspGluIle 100
301 GAGTTCGAACGCTACGCGCCGAGCGTGGCAAACTTAAGCGCGGANGA 350
|||||
101 GluPheGluArgIyrAlaProGluAlaLeuAlaAsnLeuSerGlyGluG 117
351 ANTNNNGNCAATCTGATCCAATCCGGTTTGTGGACTCGCGCTCGGTANCC 400
|||||
117 uValArgArgAsnLeuIleGlnSerGlyLeuTrpThrAlaLeuArgThrA 134
401 GTCCGTTACGAAATCCCTCGCGTCGATGCCGAGCGGTTTCGCATCTTC 450
|||||
134 rgProPheSerLysIleProAlaValAspAlaGluProPheAlaIlePhe 150
451 GTCAATCGCATGACACCAATCCGCTNGCGCGGACAGCCCTGTGTTGTAT 500
|||||
151 ValAsnAlaMetAspThrAsnProLeuAlaAlaAspProThrValIleI 167
501 CAAAGAAGCCGCGANGATTTCAGAGGANGTNTGCTGGTATTGAGCGGTT 550
|||||
167 eLysGluAlaAlaGluAspPheLysArgGlyLeuLeuValLeuSerArgL 184
551 TGACCGAGCGCTAAATCCATGTGTGAAGCAGCTGGCGCAGACGTCGCG 600
|||||
184 euThrGluArgLysIleHisValCysLysAlaAlaGlyAlaAspValPro 200
601 TCTGAAATGCTGCCAACATCGAAACACATGAATTCGGCGGCGCCCATCC 650
|||||
201 SerGluAsnAlaAlaAsnIleGluThrHisGluPheGlyGlyProHisPr 217
651 GGCCGCTTTGAGTGCGACGCACATCAATTCATTGAGCGCGTGGTCAAA 700

```

```

|||||
217 oAlaGlyLeuSerGlyThrHisIleHisPheIleGluProValGlyAlaA 234
701 ACAAACCGTTTGGACCATCAATATCAAGATGTAATTTGCCATCGGACGT 750
|||||
234 snLysThrValTrpThrIleAsnTyrglnAspValIleThrIleGlyArg 250
751 TTGTTTCCAACAGCCGCTCTGAACACCGGCGCGTGTATTGCTTTGGGTGG 800
|||||
251 LeuPheAlaThrGlyArgLeuAsnThrGluArgValIleAlaLeuGlyG 267
801 TTCTCAAGTCAACAACACCGCTCTTGGTACCGTTCGGTGGTGGCGAAAG 850
|||||
267 ySerGlnValAsnLysProArgLeuLeuArgThrValLeuGlyAlaLys 284
851 TATCGCAATTAATCTGCGGCGCAATTTGTTGACGACGACACCAACCGCGTATT 900
|||||
284 alSerGlnIleThrAlaGlyGluLeuValAspThrAspAsnArgValIle 300
901 TCCGCTTCGGTATTGAACGCGCGGATTACACAAGCGCGCGCATTTATTT 950
|||||
301 SerGlySerValLeuAsnGlyAlaIleThrGlnGlyAlaHisAspTyrLe 317
951 GGGACGCTACCAATCAGATTTCCGTTATCGAAGAGCGCGCAGCAAG 1000
|||||
317 uGlyArgTyrglnHisAsnGlnIleSerValIleGluGlyArgSerLysG 334
1001 AGCTGTTTCGGTGGTGGCGCGCGCGCGCGCAAACTTCAAGTTCACGACGC 1050
|||||
334 luLeuPheGlyTrpValAlaProGlnProAspLysTyrSerIleThrArg 350
1051 ACAGCCCTCGGCCATTCTCTGAAACAACTTCAAGTTCACGACGC 1100
|||||
351 ThrThrLeuGlyHisPheLeuLysAsnLysLeuPheLysPheAsnThrAl 367
1101 CGTCAACGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1150
|||||
367 aValAsnGlyGlyAspArgAlaMetValProIleGlyThrTyrgluArgV 384
1151 TAATGCGCGCTAGACATCTTCCCTACCGCTACCGCTTTCGGCGGATTATCGTC 1200
|||||
384 alMetProLeuAspIleLeuProThrLeuLeuLeuArgAspLeuIleVal 400
1201 GCGATACCGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1250
|||||
401 GlyAspThrAspSerAlaGlnAlaLeuGlyCysLeuLeuLeuAspGluG 417
1251 AGACCTCGCTTTGTGCGAGCTTCGCTCGCGCGCGCGCGCGCGCGCG 1300
|||||
417 uAspLeuAlaLeuCysSerPheValCysProGlyLysTyrgluTyrglyP 434
1301 CGCTGTGCGTAAGCTGCGTGGAAACCTNTGAGAGGAGGCG 1341
|||||
434 roLeuLeuArgLysValLeuGluThrIleGluLysGluGly 447

```

seq_name: pir2:164002

seq_documentation_block:

sodium-translocating NADH dehydrogenase (ubiquinone) (EC 1.6.5.-) alpha chain HI0164
 C:Species: Haemophilus influenzae
 C:Date: 18-Aug-1995 #sequence-revision 02-Sep-2000 #text_change 02-Mar-2001
 C:Accession: I64002; A64003
 R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage
 ; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman
 ; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
 Science 269, 496-512, 1995
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter
 A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A:Reference number: A64000; MUID:95350630
 A:Accession: I64002
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-447 <TIG>

1 ATGATTAAAAATCAAAAAGGCTCTAAACCTGCCCATCGCGGGCAGACCCGGA 50
1 MetIleThrLysLysGlyLeuAspLeuProIleAlaGlyLysProAl 17
51 CCAAGTCATTATCAGCGGCCCTGCATTACCAAGTCGCGTGTGTCGCG 100
17 aglnValIleHisSerGlyAsnAlaValAsnGlnValAlaIleLeuGlyG 34
101 AAGAATATCGCGGTATCGCCCTNGATGAAAGTCAAGAGGCGATGCC 150
34 luGluTyValGlyMetArgProSerMetLysValArgGluGlyAspVal 56
151 GTCAAAAAGGCCAAGTCGTGTTGAAGACAAAGNATCCGGGCGTGGT 200
51 ValLysLysGlyGlnValLeuPheGluAspLysLysAsnProGlyValI 67
201 GTTTACCGCGCCGTTTCAGGCAAAATCGCGGCATCCATCGCGCGAAA 250
67 ePheThrAlaProIleAsnSerGlyThrIleThrAlaIleAsnArgGlyGlu 84
251 AGCGCGTACTTACGTCGGTCGTGATGCGCGTTGAAGGCACACGAAATC 300
84 ysArgValLeuGlnSerValValIleAsnValGluGlyAspGluLysIle 100
301 GAGTTCGAACCTACGCGCCCGAAGCGTTGCGCAAACTTAAGCGGCGGACA 350
101 ThrPheAlaLysTySerThrGluGlnLeuAsnThrLeuLeuSerSerGluG 117
351 ANTNNGNGCAATCGATCCAAATCCGTTTGTGGACTCGCGTCGCGTANCC 400
117 nValLysGlnAsnLeuIleGluSerGlyLeuTrpThrAlaLeuArgThrI 134
401 GTCGCTTCAGCAAAATCCCTCCGTCGATCCAGCCGCGGTCCGCATCTTC 450
134 rgProPheSerLysValProSerIleGluSerGluAlaSerSerIlePhe 150
451 GTCATTCGATGGACACCAATCCGCTNGCGGCAGACCCCTGTGGTGTGAT 500
151 ValAsnAlaMetAspThrAsnProLeuAlaAspProSerValValle 167
501 CAAAGAAGCCGNGANGATTTACAGCAGANGTTCGTGGTATTGACGCCGT 550
167 uLysGluTySerGlnAspPheThrAsnGlyLeuThrValLeuSerArgL 184
551 TGACCCGACGTA . . ATCATCTGTGTAAAGCAGCTGCGCAGACAGCTG 597
184 euPheProSerLysProLeuHisLeuCysLysAlaGlyAspSerAsnIle 200
598 CGCTTCGAAAATCGTCCCAACATCGAACAACATGAATTCGGCGCCCGCA 647
201 ProThrAlaAspLeuGlnAsnLeuGlnIleHisAspPheThrGlyValHi 217
648 TCGGCGCGGTGTGACGCGACGACATTCATTTCATTGACCCGTCGGTG 697
217 sProAlaGlyLeuValGlyThrHisIleHisPheIleAspProValGlyI 234
698 CAAACAAAACCGTTGGACCATCAATATCAAGATGTAAATTCGCATCGGA 747

C; Gene: nqrA
C; Superfamily: Vibrio alginolyticus sodium-translocating NADH dehydrogenase (ubiquinol-cytochrome b₅ reductase)
C; Keywords: oxidoreductase

1

alignment_scores:

alignment_scores:

Percent Similarity: 83.893 Percent Identity: 60.403

alignment_block:

US-09-303-518D-127 x G82094

Align seg 1/1 to: G82094 from: 1 to: 462

```
1  ATGATTAAATCAAAAGGCTCTAAACCTGCCCATCGCGGCGAGACCGGA 50
|||||
17  MetileThrlleLysLysGlyLeuAspLeuProIleAlaGlyThrProse 33
|||||
51  GCAAGTCATTATGACGGCGCGCTCATTAACCGAAGTCGCTTGGCG 100
|||||
33  rGlnValIleSerAspGlyLysAlaIleLysLysValAlaLeuLeuGlyG 50
|||||
101  AGAATATATCGCGGTATGCGCCCTGATGAAAGTCAAGAAAGCGCATGCC 150
|||||
50  LuGlyThrValGlyMetArgProThrMetHisValArgValGlyAspGlu 66
|||||
151  GTCAAAAAGGCCAAGTGTCTGTTGAAGACAAAAAGNATCCGGCGTGT 200
|||||
67  ValLysLysAlaGlnIleLeuPheGluAspLysLysAsnProGlyVally 83
|||||
201  GTTTACCGCGCCNGTTTCAGGAAATCCGCCCATCCATCGCGCGGAA 250
|||||
83  sPheThrSerProValSerGlyLysValValGluIleAsnArgGlyAlaL 100
|||||
251  AGCGGTACTTCAGTCGGTGGTGTGCGGTGAAGGCAAGCAAGCAAGATC 300
|||||
100  ysArgValLeuGlnSerValValIleGluValAlaGlyAspAspGlnVal 116
|||||
301  GAGTTCGAACGCTACGCGCCCGAAGCGTTGGCAACTTAAGCGCGGANGA 350
|||||
117  ThrPheAspLysPheGluAlaAsnGlnLeuAlaSerLeuAsnArgAspAl 133
|||||
351  ANTNNNGCATCTGATCCATCGGTTTGTGACTCGCGTGGGTANCC 400
|||||
133  arileLysThrGlnLeuValGluSerGlyLeuThrAlaPheArgThrA 150
|||||
401  GTCCGTTGCAAAATCCCTGCGTGCATGCGCGAGCGCTTCCGCTATTC 450
|||||
150  rgProPheSerLysValProAlaIleAspSerThrSerGluAlaIlePhe 166
|||||
451  GTCATGCGATGGACCAATCCGCTGCGGCGAGACCGCTGTGTGTGAT 500
|||||
167  ValThrAlaMetAspThrAsnProLeuAlaAlaGluProThrValValII 183
|||||
501  CAAAGAACCCGCGANGATTTCAGACGANGTNTGCTGTATTGAGCGGTT 550
|||||
183  eAsnGluGlnSerGluAlaPheValAlaGlyLeuAspValLeuSerAlaL 200
|||||
551  TGACCGAGCGTAAATCATGTGTGTAAGCAGCTGCGCGAGAGCGCGCG 600
|||||
200  euThrThrGlyLysValThrValCysLys...LysGlyThrSerLeuPro 215
|||||
601  TCTGAAATGTCGCAACATCGAACAACATCAATTCGCGCGCGCGCATCC 650
|||||
216  ArgSerGlnProAsnValGluGluHisValPheAspGlyProHisPr 232
|||||
651  GGCGGTTTGTAGTGGCAGCACATTCATTATGAGCGCGGTGCGGCA 700
|||||
232  oAlaGlyLeuAlaGlyThrHisMetHisPheLeuThrProValSerAla 249
|||||
701  ACAAACCGTTTGCACCATCAATATCAAGTATGATTCGCTTGGCGAGCT 750
|||||
249  sPHisValAlaIleAsnThrGlnArgValValSerLeuAlaG 282
|||||
751  TTGTTTCAACAGCGCTCTGAACACGCGCTGATTCCTTGGGTGG 800
|||||
266  LeuPheLeuThrGlyGluLeuThrGlnArgValValSerLeuAlaG 282
|||||
801  TTCTCAAGTCAAAACACGCGCTTGTGCGTACCGTTTGGGTGCGAAG 850
|||||
```

```
282  yProValValAsnLysProArgLeuValArgThrValMetGlyAlaSerL 299
851  TATCGCAAATTAATCGGGCGAATTGTTGACGACGACCAACCGCGTGTAT 900
|||||
299  euGluGlnLeuValAspSerGluIleMetProGlyGluValArgIlelle 315
|||||
901  TCCGGTTCCGGTATTGACGCGCGCATTAACAAGCGCGCGCATTAATTT 950
|||||
316  SerGlySerValLeuSerGlyThrLysAlaThrGlyProHisAlaThrLe 332
|||||
951  GGGAGCGCTACCAATCAGATTTCGGTTATCGAAGAAGCGCGCGCAAG 1000
|||||
332  uGlyArgThrHisLeuGlnValSerValLeuArgGluGlyArgAspLysG 349
|||||
1001  AGCTGTTGCGGTGGTTCGCGCGCGAGCGCGCAAAATCTCCTCACGCGT 1050
|||||
349  LuLeuPheGlyTrpAlaMetProGlyLysAsnLysPheSerValThrArg 365
|||||
1051  ACGACCTCGGCCCATTTCTCTGAAACAAACTCTTCAAGTTTCACGACAG 1100
|||||
366  SerPheLeuGlyHisLeuPheLysGlyGlnValTrpAsnMetThrThrTh 382
|||||
1101  CGTCAACGGTGGCGACCGCGCATGTGCGGATTGGTACTTACGACGCG 1150
|||||
382  rThrAsnGlySerAspArgSerMetValProIleGlyAsnTrpGluLysV 399
|||||
1151  TAATGCGCGCTAGACATCCTGCTACCTGCTTTGCGCGCATTTAATCGTC 1200
|||||
399  alMetProLeuAspMetGluProThrLeuLeuLeuArgAspLeuCysAla 415
|||||
1201  GCGCATACCGACGCGCAAGCATTTGGGTGCTTGGATTTGGACGACAGA 1250
|||||
416  GlyAspSerAspSerAlaValArgLeuGlyAlaLeuGluLeuAspGluG 432
|||||
1251  AGACCTCGCTTTGTGAGCTTCGCTCGCGGCGCAAAATACGAATANGGC 1300
|||||
432  uAspLeuAlaLeuCysThrPheValCysProGlyLysTrpGlyTrpGlyG 449
|||||
1301  CGCTGTTGCTAAGTGTGCGAACCNTTGTGAGAAAGGAGGC 1341
|||||
449  InLeuLeuArgGluCysLeuAspLysIleGluLysGluGly 462
```

seq_name: pir2:H83272

seq_documentation_block:

sodium-translocating NADH dehydrogenase (ubiquinone) (EC 1.6.5.-) nqr1 chain PA2999 [C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 16-Feb-2001
R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: A82950; MUID:20437337
A:Accession: H83272
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-445 <STO>
A:Cross-references: GB:AE004724; GB:AE004091; NID:g9949083; PIDN:AAG06387.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: nqrA; PA2999
C:Superfamily: Vibrio alginolyticus sodium-translocating NADH dehydrogenase (ubiquino
C:Keywords: Oxidoreductase

alignment_scores:

Quality: 1298.00 Length: 447
Ratio: 3.576 Gaps: 1
Percent Similarity: 81.208 Percent Identity: 57.271

alignment_block:

US-09-303-518D-127 x H83272

Align seg 1/1 to: H83272 from: 1 to: 445

```

1  ATGATTAAATCAAAAAGGCTAAACCTGCCATCGCGGACAGACCGGA 50
|||||
1  MetIleLYsIleLYsArgGLyLeuAspLeuProIleSerGLyAlaProGL 17
|||||
51  GCAAGCTATTATATGACGGCGCGCTATTACCGAAGTCGCGTGTGCTGGCG 100
|||||
17  uGlnArgIleGLuAlaAlaArgProValArgSerValAlaLeuIleGLyP 34
|||||
101  AGAATATCCCGGTATCGCCCGCCCTNGATGAAGTCAAGGACGAGGATGCC 150
|||||
34  heAspTrHisGLyMetLYsProThrMetAlaValGlnValGLyAspArg 50
|||||
151  GTCAAAAAGGCGCAAGTGTGTTGAAGACAAAAAGNATCCGGCGCTGCT 200
|||||
51  ValLYsLeuGLyGlnValLeuPheThrAspLYsLYsAsnProSerValSe 67
|||||
201  GTTTTACCGCGCNGTTTTCAGGCAAAATCGCCGCCATCCATCGCGCGAAA 250
|||||
67  rTYrThrAlaProGLyAlaGLyValValSerAlaIleHisArgGLyGLu 84
|||||
251  AGCGCGTACTTCAGTCGGTCGTGATTGCGGTGAAGCGAACGACGAGAAATC 300
|||||
84  ysArgValLeuGlnSerValValIleAspLeuAspGLyAspGluGlnLeu 100
|||||
301  GAGTTCGAACGCTACGCGCCCGCAAGCGTTGGCAAACTTAAGCGCGGANGA 350
|||||
101  GluPheAlaArgTYrProAlaAspLYsLeuAlaThrLeuSerAlaGluGL 117
|||||
351  ANTNNNGNCAATCTGATCCATCCGTTGTTGGACTCGCTGCGGTANCC 400
|||||
117  nValArGaSPAsnLeuLeuGlnSerGLyLeuTrpThrAlaLeuArgThrA 134
|||||
401  GTCGGTTCAGCAAAATCCCTGCGCTGATGCGAGCGCGCTTCGCCATCTTC 450
|||||
134  rgProPheSerLYsValProAspProGluSerSerProSerSerIlePhe 150
|||||
451  GTCAAATGCGATGACACCAATCGCTNGCGCGGACAGCCCTGTGTTGTGAT 500
|||||
151  ValThrAlaIleAspThrGlnProLeuAlaAlaAspProGlnValVal 167
|||||
501  CAAAGAAGCGCGANGATTTCAGACGANGTNGCTGTTATTGAGCCGTT 550
|||||
167  eAlaGLuGlnGLyAlaAlaPheGlnAlaGLyLeuThrValLeuGLyArgL 184
|||||
551  TGACCGGAGCGTAAATTCATGTGTGAAGCGAGCTGGCGCACAGCTGGCG 600
|||||
184  euAla.....ArgValPheLeuCysLYsAlaGluGLyValSerLeuPro 198
|||||
601  TCTGAAATGCTGCCACATCGAAACACATGAATTCGGCGGCGCCGATCC 650
|||||
199  GlyGluAlaLeuSerGLyValThrAlaGlnAlaPheSerGLyProHisPr 215
|||||
651  GGCGGTTTGAGTGGCACGACATTCATTTCATTTAGCGCGGTGGTGCAA 700
|||||
215  oAlaGLyLeuProGLyThrHisIleHisPheLeuAspProValGLyAlaG 232
|||||
701  ACAAAACCGTTTGGACCATCAATATCAAGATCAATTCGCCATCGGACGT 750
|||||
232  lLYsSerValTrpAsnLeuAsnTYrGlnAspValIleAlaIleGLyLYs 248
|||||
751  TTGTTTGCACAGCGCGTCTGAACACGAGCGCGTGATTGCTTTGGGTGG 800
|||||
249  LeuPheThrThrGLyGlnLeuTYrThrGluArgValIleAlaLeuAlaGL 265
|||||
801  TTCACAGTCAACAAACACGCGCTCTTGGCTGATACGTTTGGGTGCCAAG 850
|||||
265  yProValValGluLYsProArgLeuValArgThrArgLeuGLyAlaAsnL 880
|||||
851  TATCGCAATTTACTGCGGGCGAATTGGTTTCACGACAGCAACCGCGTGAT 900
|||||
```

4 ATTTAAATCAAAAAAGGTCTAAACCTGCCCATCGCGGCAGACCGGAGCA 53

3 IleThrValAsnArgGlyLeuAspSerLeuGlnGlySerProLysG1 19
54 AGTCATTATTATGACGGCCCGTCAATACCAAGTTCGCTTGGCGAAG 103
19 uSerGlyPheTyrAsnLysIleAspProGluPheValSerIle..... 33
104 AATATGCCGGTATGGCCCGC.....TNGATGAAGATC 135
34AspLeuArgProPheGlnProLeuSerLeuLysLeuVal 47
136 AAGGAGGCGATCGCGTCAAAAAGCCAAAGTCTGTTTGAAGACAAAAA 185
48 GluGlnGlyAspAlaValCysSerGlyAlaProIleAlaGluTyrLysH1 64
186 GNATCGGCGCGTGGTGTTCACGCCCGCTTCAGSCAAATCGCCGCA 235
64 sPheProAsnThrTyrIleThrSerHisValSerGlyValValThrAlaI 81
236 TCATCTCGGCGGAAAGCGCGTACTTCAGTCGGTGGTATT...GCCGTT 282
81 leArgArgGlyAsnLysArgSerLeuLeuAspValIleIleLysLysThr 97
283 GAAGCAACAGGAGCAATCGATCCAAAGCTACCGCCCGCGAGCGTTGCC 332
98 ProGlyProThrSerThrGluTyr.....ThrTyrAspLeuG1 110
333 AAACCTAAGCGCGGANGAANTNNGCAATCTGATCAATCCGGTTGT 382
110 nThrLeuSerArgSerAspLeuSerGluIlePheLysGluAsnGlyLeuP 127
383 GGACTGCCGCTCGGTANCCGTCGTTACGAAATCCCTCGCGTGCATGCC 432
127 heAlaLeuIleLysGlnArgProPheAsp...IleProAlaIleProThr 142
433 GAG...CGTTTCGCGCTCTCGTCAATCGGATGACGACCAATCCGTCNGC 479
143 GlnThrProArgAspValPheIleAsnLeuAlaAspAsnArgProPheTh 159
480 GCGAGACCTCGTGGTTGTGATC.....AAAGAAGCCG 511
159 rProSerProGluLysHisLeuAlaLeuPheSerSerArgGluGluGlyP 176
512 NCGANGATTTCAGACGANGTGTGCTGTTGATTCAGCCGTTTCAGCAGCGT 561
176 heTyrValPheValGlyValArgAlaIleAlaLysLeuPheGlyLeu 192
562 AAAATCCATGTGTGAAGCAGCTGGCGCAGACGTCGCTGTAAGATGC 611
193 ArgProHisIleValPheArgAspArgLeuThrLeuProThrGlnGluLe 209
612 TGCCCAACATC...CAACACATGAATTCGGCGGCCCGCATCCGCGCGTT 658
209 uLysThrIleAlaHisLeuHisThrValSerGlyProPheProSerGlys 226
659 TGAGTGGCAGCAGCATTCATTTCATTCAGCGCGTTCGTCGCAACAAA... 705
226 erProSerIleHisIleHisSerValAlaProIleThrAsnGluLysGlu 242
706 ACCGTTTGGACCAATTAATCAAGATGTAATTCATTCGATCGAGGTTTGT 755
243 ValValPheThrLeuSerPheGlnAspValLeuThrIleGlyHisLeuPh 259
756 TGCAACAGCGCGTCTGAACACCGACCGCGCGTATGCTTGGTGGTGTCTC 805
259 eLeuLysGlyArgIleLeuHisGluGlnValThrAlaLeuAlaGlyThrA 276
806 AAGTCAACAAACCA.....CGCTCTTCGCTACCGGTTTGGTGGCGAAA 849
276 laLeuLysSerSerLeuArgArgTyrValIleThrThrLysGlyAlaSer 292
850 GTATCGCAA...ATTACTCGCGCGAATGTTGTTGACGACAGCAACCGCGT 896
293 PheSerSerLeuIleAsnLeuAsnAspIleSerAspAsnAspThr...Le 308

897 GATTTCGGTTTCGTTATTGAACGGCGGATTCACAAAGCGCGCAGCAT. 945
308 uIleSerGlyAspProLeuThrGlyArgLeuCysLysLysGluGluGluP 325
946 ..TATTTCGGGACGTACCAATCAGATTTCCTCGTTATCGAAGAAGCGCG 993
325 roPheLeuGlyPheArgAspHisSerIleSerValLeuHisAsnProThr 341
994 ACAGAAAGAGCTGTTTCGGCTGGCTGGCCCGCAGCGCCGACAAATCTCCAT 1043
342 LysArgGluLeuPheSerPheLeuArgIleGlyPheAsnLysProThrPh 358
1044 CACCGGTACGACCCCTCGGCCATTTCCTGAAAAACAAACTCTTCAAGTTCA 1093
358 eThrLysThrTyrLeuSerGlyPhePheLysLysArg...ThrTyrT 374
1094 CG.....ACAGCGGTCAACGGTGGCGACCGCGCCATGGTGGCGATT 1134
374 hrAsnProAspThrAsnLeuHisGlyGluThrArgProIleIleAspThr 390
1135 GGTACTTTACGAGCGGTAAATCGCGCTAGACATCTGCTACCCCTGCTTTT 1184
391 AspIleTyrAspLysValMetProMetArgIleProValValProLeuI 407
1185 GCGCGATTAAATCGTCGGCGATACCGACGCGCAAGCATTTGGTGTGT 1234
407 eLysAlaValIleThrLysAsnPheAspLeuAlaAsnGluLeuGlyPheL 424
1235 TGGAAATGGACGAAGACCTCGCTTGTGCGAGCTTCGTCGCCCGCGGC 1284
424 euGluValCysGlyGluAspPheAlaLeuProThrLeuIleAspProSer 440
1285 AAATACAGATANGCGCGCTGTTGCTGAAGTCTCGTGAACCCNTTGAGAA 1334
441 LysThrGluMetLeuThrIleValLysGluSerLeuIleGluTyrAlaL 457
1335 GGAA 1338
457 sGlu 458

seq_name: p1r2:E72040

seq_documentation_block:

probable sodium-translocating NADH dehydrogenase (ubiquinone) (EC 1.6.5.-) alpha chain
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 16-Feb-2001
C:Accession: E72040; G81623
R:Kallman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, N. Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606
A:Accession: E72040
A:Molecule type: DNA
A:Residues: 1-467 <ARN>
A:Cross-references: GB:AB001656; GB:AB001363; NID:g4377047; PIDN:RAD18882.1; PID:g437
A:Experimental source: strain CWL029
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Complete sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39
A:Reference number: A81500; MUID:20150255
A:Accession: G81623
A:Molecule type: DNA
A:Residues: 1-467 <REA>
A:Cross-references: GB:AE002164; GB:AE002161; NID:g718939; PIDN:AAF37899.1; PID:g718
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: nqrA; CP0002
C:Superfamily: Vibrio alginolyticus sodium-translocating NADH dehydrogenase (ubiquino
C:Keywords: oxidoreductase

A:Accession: F71489
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-465 <ARN>
A:Cross-references: GB:AE001334; GB:AE001273; NID:g3329078; PIDN:AAE68238.1; PID:g332
A:Experimental source: Serotype D, strain UW-3/Cx
C:Genetics:
C:Gene: ngrA
C:Superfamily: Vibrio alginolyticus sodium-translocating NADH dehydrogenase (ubiquinol)

A;Residues: 1-1802 <DIE>

A;Residues: 1-1802 <DIE>

R.Kasahara, S.; Yamada, H.; Mio, T.; Shiratori, Y.; Miyamoto, C.; Yabe, T.; Nakajima, T.
J. Bacteriol. 176, 1488-1499, 1994
A:Title: Cloning of the *Saccharomyces cerevisiae* gene whose overexpression overcomes the
A:Reference number: A53382; MUID:94156857
A:Accession: A53382

A:Molecule type: DNA
A:Residues: 1-581, 'A', 583-593, 'A', 595-1802 <KAS>
A:Cross-references: EMBL:S69101; NID:9545659; PIDN:AAB30051.1; PID:9545660
A:Experimental source: YNN295
A:Note: sequence extracted from NCBI backbone (NCBIN:144410, NCBIIP:144411)
C:Genetics:

A:Gene: SGD:HKR1
A:Cross-references: SGD:S0002828; MIPS:YDR420w
A:Map position: 4R

C:Keywords: calcium binding; glycoprotein; transmembrane protein
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-1802/Product: HKR1 protein #status predicted <NAT>
F:1483-1508/Domain: transmembrane #status predicted <TM>
F:1645-1656/Domain: calcium binding #status predicted <CAL>

alignment_scores:
Quality: 133.50 Length: 445
Ratio: 0.588 Gaps: 20
Percent Similarity: 51.011 Percent Identity: 22.921

alignment_block:

US-09-303-518d-127/rev x S69703

Align seg 1/1 to: S69703 from: 1 to: 1802

```
1337 TCCCTCTCAAGGTTTCCACACCTTACGCAACAGCGGCGCCTATTCGTA 1288
|||||
390 SerAspThrValValSerThrSerThrAsnThrValProTyrSerSe 406
1287 TTTCGCCGGCAGACGAGTGTGCAAGCGAGGTCTTCTTCGTCGAAT 1238
|||||
406 rVal...HisSerThrPheValHis...AlaThrSerSerThrTyrI 421
1237 CCAAGCAACCAATGCTTGGCGCTGTCGATGCGGAGGATTAATCG 1188
|||||
421 leSerSerSerLeuTyrSerSerProSerLeuSerAlaSerValSer 437
1187 CGCAAAAGCAGGTPAGGAGGATGTAGCGGATTACGCGCTCGTAAGT 1138
|||||
438 HisPheGlyValAlaProPheProSerAlaTyrIleSerPheSerVa 454
1137 ACCAATC.....GGCACCATGGCGGTCGCCACCGTTGACGCGTG 1097
|||||
454 lProValAlaValSerSerThrTyrThrSerSerProSerAlaSerVal 471
1096 TCGTGAACCTTGAAGAGTTGTTTTCAGAAATGGCCGAGGTCGTACGC 1047
|||||
471 alVal..... 472
1046 GTGATGGAGTATTGTCGGCTGGCGGCACCCAGCGCAACAGCTCTTT 997
|||||
473 .....ProSerAlaTyrAlaSerSerProSerValProVa 484
996 GTCGGCGCTTTCGATAACGGAATCTGATTGTTGGTAGCTCCCAAT 947
|||||
484 lAlaValSerSerThrTyrThrSerSer..... 493
946 AATCGTGGCGCTGTGTGTAATC...GGCCGTTTCAATACGAAACCGGAA 900
|||||
494 ..ProSerAlaProAlaAlaSerSerThrTyrThrSerSerProSer 509
899 ATC.....ACCGGTTGTCTCGTCAACCAATTCGCCGCGC 865
510 AlaProValAlaValSerSerThrTyrThrSerProSerAlaProAl 526
864 AGTAATTTGCGATACTTTC.....GCACCCAAACGGGTAC 830
|||||
```

seq_name: pir2:AE0692

seq_documentation_block:

probable NADH reducing dehydrogenase STY1665 [imported] - Salmonella enterica subsp.
C:Species: Salmonella enterica subsp. enterica serovar typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001

```
526 aAlaIleSerSerThrTyrThrSerSerProSerAlaProValAlaVal. 542
829 CCAAGAGCGGTGTTTGTGACTTGAACAACA.....CCCAAGCAATC 786
543 .....SerSerThrTyrThrSerSerProSerAlaProAlaAlaIle 556
785 ACGCGCTCGGTGTTTCAGACGGCTGTTGCAACAACGTCCTCGATGCAAT 736
557 SerSerThrTyrThrSerSerProSerAla.....ProValAlaVa 570
735 TACATCTTGATAATTGATGTTCCAAACGGTTTCTTGTTCACCGACCGCT 686
570 lSerSer.....ThrTyrThrSerSerProSerAla 581
685 CAATGAATGAATGTGCTGCCACTCAACACGGCGGATGCGGCGCGCG 636
581 ro.....ValAlaIleSerSerThrTyrThrSerSerPro 592
635 AATTCATCTGTTTCGATGTTGGCAGCATTTTCAGACGGCAGCTCTGCGC 586
593 SerValProValAlaValSerSerThrTyrThrSerProSerAlaPr 609
585 AGCTGCCCTTACACACATGGATTTTACGCTCGTCAAA.....CGGC 545
609 oAlaAlaIleSerSerThrTyrThrSerSerProSerAlaProValAla 626
544 TCAATACCAGCANACNTCGTCTGAATCNTCGNCGGCTTCTTTGATCACA 495
626 alSerSerThrTyrThrSerSerProSerAlaProAlaAlaIleSer 642
494 ACCACAGGCTGCGCGNAGCGGATTTGTTCCATCCGATTCGACGAGAT 445
643 ThrTyrThrSerSerProSerValProValAlaValSerSerThrTyr 659
444 GCGCAACGGCTCGCATCGACGGCAGGATTTTCTGAACGGCAGCGGTAC 395
659 rSerSerProSerAlaProAlaAla.....IleSerSerThrTyrT 673
394 GCGCGCAGTCCCAACACCGGATTTGATCAGATTCGNCNNNANTTCNCG 345
673 hrSerSerProSerValProValAlaValSerSerThrTyrThrSer 689
344 CGCTTAAGTTTCCCAACGGCTGCGGCGGTAGCTTCGAACTCGATTC 295
690 Pro.....SerAlaProAlaAlaIleSerSerThrTyrTh 701
294 GTCTGTTGCTTCAACGGCAATCACGACCGACTGAAGTACGCGCTTTTCG 245
701 rSerSerProSerAlaProValAlaValSerSerThrTyrThrSerSer 718
244 CGCGATGGATGGCGGATTTTCCCTGAACNGCGCGGTAAACACACCG 195
718 roSerAlaProAlaAlaIle.....SerSerThrTyrThrSerSer 731
194 CCGGATNCTTTTGTCTCAACACGACTTGGCTTTTTCGACGGCATC 145
732 ProSerAlaProValAlaValSerSerThrTyr.....ThrSer 745
144 GCCTCTCTTGACTTTTCATCAGGGCGCATACCGGCA..... 108
745 rProSer.....AlaProAlaAlaIleSerSer 755
107 .....TATTCTTCGCCCAACGACGCTTCGTA 78
755 hrTyrThrSerSerProSerAlaProValAlaVal 766
```

572 TGTGAAGGCAGCTGGCGCAGACGTGCCGCTGAAATGCTGCCAACATC 621

R. Churchel, C.M.
submitted to the EMBL Data Library, September 1995

A:Reference number: S59302
A:Accession: S59310
A:Molecule type: DNA
A:Residues: 1-1104 <CHU>
A:Cross-references: EMBL:Z54141; GSPDB:GN00013; MIPS:YMR317w
A:Experimental source: strain AB972
C:Genetics:
A:Gene: MIPS:YMR317w
A:Map position: 13R

alignment_scores:
Quality: 131.00 Length: 466
Ratio: 0.577 Gaps: 16
Percent Similarity: 48.712 Percent Identity: 22.532

alignment_block:

US-09-303-518D-127/rev x S59310 ..

Align seg 1/1 to: S59310 from: 1 to: 1104

```
1331 TCANGGTTTCCACACCTTACGCAACAGCGGCGCCTATTCGTATTGCC 1282
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
350 SerSerThrSerSerValSerSerGluAlaProSerSer..... 363
1281 CGGCAGACGAAGCTGCACAAAGCGAGGTCTCTCTCGTCCAAATCCCAAGC 1232
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
364 .....ThrSerSerValSerSerGluA 372
1231 AACCAATGCTTGGCGGTGTCGATCGCGAGCATTAATCGCGCAAA 1182
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
372 laProSerSerThrSerSerValSerSerGluAlaProSerSerThrLys 388
1181 AGCAGGTAGGACAGTGTAGCGGATTCAGCGCTCGTAAGTACCAAT 1132
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
389 SerSerVal.....MetSerSerGluValSerSerAla.....Th 400
1131 CGGCACCATGCGCGGTGCGCCACCGTTCAGCGCTGCTGTAACCTGAAGA 1082
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
400 rSerSerValSerSerGluAlaProSerAlaIleSerSerLeuAlaS 417
1081 GT.....TTGTTTTCAGGAATGGCGAGGTCGTACGCGTGATG... 1041
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
417 erSerArgLeuPheSerSerLysAsnThrSerValThrSerSerLeuVal 433
1040 .....GAGTATTGTCGGGTGCGGCCACCGCACCGACACAGCTCTTT 997
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
434 AlaThrGluAlaSerSerValThrSerSerLeuArgProSerSerGluTh 450
996 GTCGGCGCTTCTTCGATAACGGAATCTGATTGTTGTTAGCTCCCAAT 947
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
450 rLeuAlaSerAsnSerIleIleGlu..... 458
946 AATCGTGGCGCCTTGTGTAATGCGCGCTTCAATACCGAACCGGAATC 897
458 ..... 458
896 AGCGGTTGTCGTCAACCAATTCGCCCGCAGTAATTCGGATCTTT 847
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
459 SerSerLeuSerThrGlyThrAsnSerThrValSerThrThrThrSerAl 475
846 CGCACCAACAGGTACGCAAGCGGTGTTGTTGTTGTTGTTGTTGTTGTTGTT 797
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
475 aAlaSerThrLeuGlySerLys..... 483
796 CCAAGCAATACGCGCTCGGTGTTTCAGACGCGCTGTTGCAACCAACAGT 747
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
484 .....ValSerSerAsnSerArgMetAlaThrSerLys... 495
746 CCGATGGCAATACATCTGTGTAATGATGTTCCAAACGGTTTCTTT... 699
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
496 ...ThrSerSerThrSerSerAspLeuSerLysSerSerValIlePheGl 511
```

```
698 .....GCACGACCGCGCTCAATGAATGAATGATGT 671
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
511 yAsnSerSerThrValThrThrSerProSerAlaSerIleSerLeuThrA 528
670 GGTGCCCTCAACCGCGCGGATGCGCGCGCGAATTCATGTGTTTCG 621
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
528 laseProLeuProSerValTrpSerAspIleThrSerSerGluAlaSer 544
620 ATGTTCGCGAGCATTTTCACGCGGACGCTCGCGCAGCTGCTTACACAC 571
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
545 SerIleSerSerAsnLeuAlaSerSerAlaProSerAspAsnAsnSe 561
570 ATGGATT.....TTACGCTCGGTCAACGCGCTCAATACCAACA 533
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
561 rThrIleAlaSerAlaSerLeuIleValThrLysThrLysAsnSerValV 578
532 NACNTCGTCTGAAATCNCGCGGCTTCTTGTATCACAACACACAGGCT 483
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
578 alSerSerIleValSerSerIleThrSerSerGluThrThrAsnGluSer 594
482 ...GCCGACGCGGATTTGCTCCATCGCATTCGCAAGAGTGGCGACGG 436
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
595 AsnLeuAlaThrSerSerThrSerLeuLeuSerAsnLysAlaThrAlaA 611
435 CTCGCGATCGACGCGAGGATTTGCTGAACGCGCGGNTACGACCGCAG 386
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
611 gSerLeuSerThrSer.....AsnAlaThrSerAlaSerAsnV 624
385 TCACAAACCGGATTCGATCAGATTCNNNNNTTCNTCGCGCTTAAG 336
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
624 alProThrGlyThrPheSerSerMetSerSerHisThrSerValIleThr 640
335 .....TTTCCCAACGCTTCGGCGCGTAGCGTTCGAACTCGATT..... 297
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
641 ProGlyPheSerThrSerSerAlaSerLeuAlaIleAsnSerThrValVa 657
296 ....TCGCTGTCCTTCAACGCGCAATCAGCAGCAGTCAAGTACGCGCT 251
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
657 lSerSerSerLeuAlaGlyThrSerPheSerThrProGluSerSerProT 674
250 TTTCCGCGGATGGATGGCGGATTTTCCTGAAACGCGCGCTAAAC 201
674 hrThrSerThrLeuValThrSerGluAlaProSerThrValSerSerMet 690
200 ACCACGCGCGATTCCTTTTGTCT.....TCAA 172
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
691 ThrThrSerAlaProPheIleAsnAsnSerThrSerAlaArgProSerPr 707
171 CAGCACTTGGCTTTTTCGCGCATCGCTTCC..... 138
707 oSerThrAlaSerPheIleThrGluSerThrSerIleSerSerValP 724
137 .....TTGACTTTTCATCAGGCGCATACCG 111
724 roLeuAlaSerGlyAspValThrSerSerLeuAlaAlaHisAsnLeuThr 740
110 GCATATTCCTCCCAACGCGCATTCGGTAATGACGCGCGCTCA 63
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
741 ThrPheSerAlaProSerThrSerSerAlaGlnLeuValSerLysSer 756
```

seq_name: pir2:I47141

seq_documentation_block:

gastric mucin (clone PGM-2A) - pig (fragment)

C:Species: Sus scrofa domestica (domestic pig)

C>Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 03-Nov-2000

C:Accession: I47141; S55315

R:Turner, B.S.; Bhaskar, K.R.; Hadzopoulos-Cladaras, M.; Specian, R.D.; Lamont, J.T.

A:Title: Pig gastric mucin: isolation and characterization of a cDNA clone with a nov

A:Reference number: I47141; MUID:94102478

A:Accession: I47141

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
 A:Residues: 1-2232 <GSI>
 A:Cross-references: EMBL:U08046; PIDN:AAC70890.1; GSPDB:GN00028; CESP:K06A9.1a
 A:Experimental source: strain Bristol N2; clone K06A9
 C:Gene: CESP:K06A9.1a
 A:Map position: X
 A:Introns: 38/1; 75/3; 103/3; 132/2; 158/2; 222/1; 1088/1; 1367/1; 2039/1; 2049/1; 2075/1

alignment_scores:
 Quality: 124.50 Length: 485
 Ratio: 0.571 Gaps: 18
 Percent Similarity: 44.948 Percent Identity: 22.062

alignment_block:

US-09-303-518D-127/rev x T34434 ..

Align seg 1/1 to: T34434 from: 1 to: 2232

```

1340 CTTCTCTCTCAANGTTTCCAGCACCTTACGCAACAGC.....GGGCC 1297
||||| |||.....: ||| |||
670 ProSerGlySerLeuGlyThrGlnSerThrAsnSerProSerPr 686
||||| |||.....: ||| |||
1296 NTATTCCTATTGCGCGGAGCAGCAAGCTGCACAAAGCGAGGTCTCTT 1247
||| |||.....: ||| |||
586 oSerLeuSerProSerThrSerGlyMetSerThrLeuThrSerGlup 703
||||| |||.....: ||| |||
1246 CGTCAATTCACCAACCAAGCTTGGCGGTGTCGGTA..... 1206
||| |||.....: ||| |||
703 roSerProSerThrGlnSerSerGlyAlaGlnSerThrLeuThr 719
||||| |||.....: ||| |||
1205 ...TCGCGGAGGATTAATCGGCAAAAGCAGGCTAGGAGGTCTAG 1159
||||| |||.....: ||| |||
720 ProSerProAsnProSerGlnSerThrSerSerLeuGluSerThrSe 736
||||| |||.....: ||| |||
1158 CGGATTACGCGCTCGTAAGTACCAATCGGCACCATGGCGGTCGCCAC 1109
||||| |||.....: ||| |||
736 rGlyAlaThrThrSerSerGlySerAlaGlyThrMetThrSerPro 753
||||| |||.....: ||| |||
1108 GCTTGACGCGTGTGTAACGTTGAAGAGTTGTTTTCAGCAATGCGCG 1059
||||| |||.....: ||| |||
753 erGlnSerSerValGly..... 759
1058 AGGTCGTACGCGTGATGGAGTATTGTCGGCTGCGGCAACCCAGCC 1009
||||| |||.....: ||| |||
760 .....SerSerGlnGlySerThrSerPr 767
1008 GAACAGCTCTTTGTCGGGCTTCTTCGATAACGGAATCTGATGTGGT 959
||| |||.....: ||| |||
767 oAlaAlaSer..... 770
958 AGCGTCCCAATATATCGTGGCGCTTGTGTAATCGCGCGCTTCAATACC 909
||||| |||.....: ||| |||
771 .....ThrThr 772
908 GAACCGGAATACACGGTGTCTGCGTCAACCAATTCGCCCGCAGTAAT 859
||| |||.....: ||| |||
773 SerGlyGluMetThrSer...GlnGlySerThrGlnThrProGlySerSe 788
||||| |||.....: ||| |||
858 TTGCATCTCTTCGACCCCAACAGGTACGCAAGAGCGGTGTTGTGA 809
||| |||.....: ||| |||
788 rValSerThrSerAlaAlaLeuLeuThrSerThrGlnGlnSerValSer 805
||| |||.....: ||| |||
808 CTTGAGAACCCCAACCAATACAGCGCTCGGTGTTTCAGACGCGCTGT 759
||| |||.....: ||| |||
805 hrAsnSerProGlySerThrValThr..... 813
758 GCAACAAACACCGCATGGCAATTAATTCGATAATTTGATGGTCCAAAC 709
||||| |||.....: ||| |||
814 .....ArgProSerThrValSerGlySerThrSerSerGlySerTh 827
708 GGTGTTGTTGACCGCACC.....GCCTCAATGAATGAA 674

```

```

827 rValThrValGlySerThrGluAlaSerThrSerGlySerValAla 844
||| |||.....: ||| |||
673 TGTGCTGCCACTCAACCGCGCGGATGCGCGCGCGAATTCATGTGTT 624
||| |||.....: ||| |||
844 erSerSerPro...AlaProSerThrSerGlnAsnProAsnProSerThr 859
||| |||.....: ||| |||
623 TCGATGTTGGCAGCATTT..... 606
||| |||.....: ||| |||
860 SerSerGlySerSerMetIleThrGlnSerProTyrProSerGlnSerTh 876
||||| |||.....: ||| |||
605 .....TCAGACGCGCAGCTGCGCCAGCTGCCCTTACACACAT 569
||||| |||.....: ||| |||
876 rSerProValGluSerSerThrProSerProGlySerProGlyThr 893
||||| |||.....: ||| |||
568 GGATTTTACGCTCGTCAACCGCTCAATACGACANACNCTCGTCTGAAA 519
||| |||.....: ||| |||
893 hrLeu.....ThrSerThrSerProSerPro 901
||||| |||.....: ||| |||
518 TCNTCGNCGGCTCTTTGATCACACACAGGGTCTGCCGCGNAGGGATT 469
||||| |||.....: ||| |||
902 SerGlnSerThrThrIleGlySerThrGlnGlySerThrSerProGly 918
||||| |||.....: ||| |||
468 GGTGTCCATCCGATTGACGAGATGCGCAAGCGCTCGCATCGACGCGCAG 419
||| |||.....: ||| |||
918 eSerThrThrSerGluMetThrSerGlnGlySerThrGlnThrPro. 934
||| |||.....: ||| |||
418 GGATTTTGTGTAACGCGAGCTACGCGAGTCCACAAACCGGATTGG 369
||| |||.....: ||| |||
935 .....GlySerThrGlySerThrValThrGlnProSerThr 946
||||| |||.....: ||| |||
368 ATCAGATGTCNNNNANTCN...TCGCGGCTTAAGTTTGGCAACGCTTC 322
||| |||.....: ||| |||
947 ValSerAspSerThrSerSerGlySerThrValThrValGlySerThr 963
||| |||.....: ||| |||
321 GGGCGGTAGCGTTCGAACTCGATTTCGCTTCCGCTTCAACGCGCAATCA 272
||| |||.....: ||| |||
963 uGly.....SerSerSerProIleProSerThrSerGlnA 975
||||| |||.....: ||| |||
271 CGACCGACTGAAGTACGCGCTTTTCGCGCGGATGGATGGCGGCGATTGG 222
||| |||.....: ||| |||
975 snThrAsnProSerThrSer...SerGlySerSerMetSerThrGlnThr 990
||||| |||.....: ||| |||
221 CCTGAACGCGC.....GCGGTAACACACGCGCGGATN 187
||||| |||.....: ||| |||
991 ProGlnSerSerGlnSerThrSerProValGluSerSerThrSerGlyAl 1007
||||| |||.....: ||| |||
186 CTTTTTGTCTTCAACACAGCATTGGCTTTTGTGACGCA.....T 146
||| |||.....: ||| |||
1007 aThrSerSerSerGlySerProGlyThrThrLeuThrSerIleSerProS 1024
||||| |||.....: ||| |||
145 CGCTTCTCTGACTTTCATCNAGGGGCGCATACCGGCATATCTTCGCCA 96
||||| |||.....: ||| |||
1024 erProSerProSerSerThrIleGlySerSerGlnGlySerThrSerPro 1040
||||| |||.....: ||| |||
95 AGCAACGCGGACT..... 84
||||| |||.....: ||| |||
1041 valValSerThrIleSerGlnGlySerThrGluThrProGlySerThrGl 1057
||||| |||.....: ||| |||
83 ..TCGGTAATGACGCGCGCTTCATAAATGACTTCTCCGCTCGCCGCCGA 35
||||| |||.....: ||| |||
1057 ySerThrValThrLysProSer.....ThrValSerGlySerAlaSerS 1072
||| |||.....: ||| |||
34 TGGGC 30
|||
1072 erGly 1073

```

seq_name: pir2:A40670

seq_documentation_block:

nuclear envelope protein POM 121 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 05-Nov-1999

C;Accession: A40670
R;Hallberg, E.; Wozniak, R.W.; Biobel, G.
J. Cell Biol. 122, 513-521, 1993
A;Title: An integral membrane protein of the pore membrane domain of the nuclear envelope
A;Reference number: A40670; MUID:93328754
A;Accession: A40670
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1199 <HAL>
A;Cross-references: GB:221513; NID:q396746; PIDN:CAA79725.1; PID:q396747
F:803-807,845-849,956-960,1010-1014,1047-1051,1076-1080/Region: pentapeptide motif (X-F-
alignment_scores:
Quality: 123.50 Length: 468
Ratio: 0.542 Gaps: 17
Percent Similarity: 48.718 Percent Identity: 20.513
alignment_block:
US-09-303-518D-127/rev x A40670 ..

Align seg 1/1 to: A40670 from: 1 to: 1199

```
1340 CTTTCCTCTCAANGTTTCAGACACTTACGCAACAGCGGCGCNATTC 1291
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
546 ProProPheThrLeuThrLeuProThrValGlyProAlaAlaSerProAl 562
1290 GTATTTGCCGGGCGAGACG.....AAGCTGCACAAAG 1259
562 aserLeuProAlaProSerSerAsnProLeuLeuGluSerLeuLysM 579
1258 CGAGCTCTTCT...TCGTCAATTCAGCAACCAATGTTGCGCGCTG 1212
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
579 etGlnGluSerProAlaProSerSerSerGluProProGluAlaAlaThr 595
1211 TCGGTATCGCGGAGGATTAAATCGCAAGAGGAGGTAGCGAGGATCT 1162
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
596 ValAlaAlaProSerProProLysThrProSerLeuLeuAlaProLeuVa 612
1161 TAGCGGATTACGCGCTCGTAAGTACCAATC...GGCACCATGGCGGGT 1115
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
612 lSerProLeuThrGly.....ProLeuAlaSerThrSerSerAspS 626
1114 CGCCACCGTTGACGGCTGCTGTAAGTGAAGT.....1080
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
626 erLysProThrThrThrPheLeuGlyLeuAlaSerAlaSerAlaThr 642
1079 ...TTGTTTTTCAGGAATCGCGAGGTCGTACGCGTGTAGTATTT 1033
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
643 ProLeuThrAspThrLysAlaProGlyValSerGlnAlaGlnLeuCysVa 659
1032 GTCGGGTGCGGCGCAACCGCGGCTCT.....999
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
659 lSerThrProAlaAlaThrAlaProSerProThrProAlaSerThrLeuP 676
998 .....TTGCTGGCGCTCTTCGATAACGGAATCTGATGTGGTAG 957
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
676 heGlyMetLeuSerProProAlaSerSerSerSerLeu.....Ala 689
956 CGTCCCAATAATCGTGGCGGCTTGTGTAATCGCGCGCTC.....915
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
690 ThrProGlyProAlaCysAlaSerProMetPheLysProLysPheProAl 706
914 .....AATACCGAACCGGAATTCACGCGGTGTCTCGCT 881
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
706 aThrProLysSerGluSerAspAsnProLeuProThrSerSerAlaAla 723
880 CAACCAATCGCGCGAGTAATTGGCATACTTCCACCCCAACCGGTA 831
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
723 laThrThrThrProAla.....SerThrAlaLeuProThrThrAla 736
830 CGCAAGAGCGGTGTTGTTGACTTGAGAACCAACCAACCAACCAACG 781
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
737 .....ThrAlaThrAl 740
780 CTCGGGTTCAGAGCGCTGTTGCAACAAACAGTCGATGGCAATTACAT 731
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
740 aHisThrPheLysProLysPheGluSerValGluProPheAla.....754
730 CTTGATAATTGATGGTCCAAACGCTTTGTTTCACCGGCGGCTCATG 681
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
755 .....AlaMetProLeuSerProProPheSerLeu 764
680 AATGATGTCGTCGCTCAACACCGCGGATGGCGCGCGCAATTC 631
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
765 LysGlnThrThrAlaProAlaThrThrAlaAlaThrSerAlaPro.....779
630 ATGTGTTTCGATTTGGCAGCATTTTCAGACGCGCATCTGCGCCAGT 581
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
780 .....LeuLeuThrGlyLeuGlyThrAlaThrSerThrValAla 793
580 CCTTACACACATGGATTTTACGCTCGTCAACAGGCTCAATACCAGCANA 531
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
793 hr.....GlyThrThrAla 797
530 CNTGCTCTGAATATCNCNGCGCTCTTTGATCACAACACAGGCTCTGC 481
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
798 SerAlaSerLysProValPheGlyPheGlyValThrThrAlaAlaSerTh 814
480 CGCNAGCGGATTTGGTCTCCATTCGATTCAGAGATGGGGAACGCTCGG 431
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
814 rAlaSerThrIleAlaSerThrSerGlnSerIleLeuPheGlyGlyAla 831
430 CATCGACGCGGAGGATTTTGTGTAACGCGGNTACGCGCGCAGCTCCAC 381
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
831 roProValThr.....AlaSerSerSer 838
380 AACCGGATTTGATCAGATTCGTCGNNNNNTTCGCGCGCTT...AAGTT 334
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
839 AlaProAlaAlaAlaSerIlePheGlnPheGlyLysProLeuAlaProAl 855
333 TCGCAACGCTTCGGCGGCTACGCTTCGAACTCGATTTTCGCTGCGCTT 284
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
855 aAlaSerValAlaGlyThrSerPheSerGlnSerLeuAlaSerSerAla 872
283 CAACGCGCAATCACGACGCTGAAGTACGCGCTTTTCGCGCGGATGATG 234
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
872 lnThrAlaAlaSerAsn.....877
233 GCGGCGATTTTCCTGTAACNCGCGGTAACACCGCGCGGATNCTT 184
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
878 .....SerSerGlyGlyPheSerGlyPheGlyThrLe 889
183 TTTGCTTCAACAGCACTTGGCTTTTTCGCGCATATCTTCGCAACGCGG 134
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
889 uThrThrSerThrSerAla...ProAlaThrThrSerGlnProThrLeu 905
133 CTTTCATCAGGGCGCATACCGCATATCTTCGCAACGCGGACT 84
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
905 hrPheSerAsnThrValThrProThrPheAsnIleProPheSerAlaSer 921
83 TCGGTAAATGACGGCGCGCTCATTAATGACTTGCCTGCGCGGCTGCGCG 34
||||| : : : : : : : : : : : : : : : : : : : : : : : : :
922 AlaLysProAlaLeuProThrTyrProGlyAlaAsnSerGlnProThrPh 938
33 GGGC 30
|||||
938 egly 939
```

seq_name: pirl:S48478

seq_documentation_block:

glucan 1,4-alpha-glucosidase (EC 3.2.1.3) - yeast (Saccharomyces cerevisiae)
N:Alternate names: extracellular glucosylase; mucin-like protein MUC1; protein Y1R01
C:Species: Saccharomyces cerevisiae
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 12-Nov-1999

C:\Accession: S48478; A26877; B26877; S27281; JC6123
R:Rowley, K.
submitted to the EMBL Data Library, October 1994
A:Reference number: S48478
A:Accession: S48478
A:Molecule type: DNA
A:Residues: 1-1367 <ROW>
A:Cross-references: GB:Z47047; EMBL:Z38061; NID:g603997; PID:g763364; GSPDB:GN00009; MIPS:
J:Yamashita, I.; Nakamura, M.; Fukui, S.
R:Bacteriol. 169, 2142-2149, 1987
A:title: Gene fusion is a possible mechanism underlying the evolution of STA1.
A:Reference number: A91831; MUID:87194600
A:Accession: A26877
A:Molecule type: DNA
A:Residues: 1-242 <YAM>
A:Cross-references: EMBL:M16164; NID:g172522; PIDN:AAA35014.1; PID:g172525
A:Accession: B26877
A:Molecule type: DNA
A:Residues: 762-1331 <YA2>
A:Cross-references: EMBL:M16165; NID:g172523; PIDN:AAA35015.1; PID:g172526
R:Pardo, J.M.; Tanex, E.; Zalacain, M.; Claros, M.G.; Jimenez, A.
FEBS Lett. 239, 179-184, 1988
A:title: Similar short elements in the 5' regions of the STA2 and SGA genes from Saccharomyces cerevisiae.
A:Reference number: S27281; MUID:89031230
A:Accession: S27281
A:Molecule type: DNA
A:Residues: 1-31 <PAR>
A:Cross-references: EMBL:X13857; NID:g4551; PIDN:CAA32069.1; PID:g4552
R:Lambrechts, M.G.; Bauer, F.F.; Marmur, J.; Pretorius, I.S.
Proc. Natl. Acad. Sci. U.S.A. 93, 8419-8424, 1996
A:title: Nuc1, a nuclein-like protein that is regulated by Mss10, is critical for pseudohyphal growth in yeast.
A:Reference number: JC6123; MUID:96332327
A:Accession: JC6123
A>Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-1367 <LAM>
A:Cross-references: GB:U30626; NID:g1304386; PIDN:AAC49609.1; PID:g1304387
C:Genetics:
A:Gene: SGD:MOC1; STA2; MAL5; DEX2; SGD:S0001458
A:Cross-references: MIPS:YIR019c; SGD:S0001458
A:Map position: 9R
C:Superfamily: yeast glucan 1,4-alpha-glucosidase homolog; glucan 1,4-alpha-glucosidase
C:Keywords: glycosidase; hydrolase; polysaccharide degradation; transmembrane protein
F:5-21/Domain: transmembrane #status predicted <TM1>
F:1350-1366/Domain: transmembrane #status predicted <TM2>

alignment_scores:
Quality: 123.00 Length: 532
Ratio: 0.498 Gaps: 23
Percent Similarity: 46.429 Percent Identity: 20.489

alignment_block:
US-09-303-518D-127/rev x S48478 ..

Align seg 1/1 to: S48478 from: 1 to: 1367

1337 TCCTTCTCAANGTTTCAGCACCTTAGCAACAGCGGCCCATTTTCGTA 1288
||||||| : ::::: |||
891 SerPseSerThrGlyThrValThrProSerSerLysTyr..... 905

1287 TTTCGGCCGGCAGACGAACATGCACAACAGGAGGTCTTCTCGTCCAATT 1238
||||||| : ::::: |||
906 ...ProGlySerGlnThrGluThrSerValSerSerThrThrGluThrf 921
||||||| : ::::: |||
1237 CCAAGCAACCCAAATGCTTCGGCGCTCGGTA.....TCGCGG 1200
||||||| : ::::: |||
921 hrIleValProThrLysThrThrSerValThrThrProSerThrThr 937

1199 AGATTAAATTCGCGAAAAAGCAGGGTAGCAGATGTCTAGCGGCATTC 1150
||||||| : ::::: |||
938 ThrIleThrThrValCvsSerThrGlyThrAsnSerAlaGlyGluTh 954
||||||| : ::::: |||


```
512 ....NCGGCTCTTTGATACAAACACAGGCTCTGCGCGNACGGATTGG 467
3158 rProThrAlaThrMetSerThrAlaThrProSerSerThrProGluThrV 3175
466 TG.....TCCATCGCATTGACGAGATGGCGAAGCGGTTCGGCATCGAGC 423
3175 aHisThrSerThrValLeuThrThrAlaThrThrThrGlyAlaThr 3191
422 GCAGGATTTCGTGAACGAGGNTACGACGCGAGTCCACAAA..... 378
3192 GlySerValAlaThrProSerSerThrProGlyThrAlaHisThrThrLy 3208
377 ....CCGATTGATGATTCGNNCNCNNANTCNCGCCGCTTAAGTTTG 332
3208 sValProThrThrThrThrGlyPheThrAlaThrPro..... 3221
331 CCAACGCTTCGGCGGTAGCTGCGTTCGAACTCGATTTTCGTTCGCTTCA 282
3222 .....SerSerProGly 3226
281 ACGGCAATCAGACGACTAGTACGCGTTTTCGCCGCGATGATGCG 232
3227 ThrAlaLeuThr.....ProProValTrpIleSe 3236
231 GCGGATTTTCCTGAACGCGCGGTAAACACACGCGCGGATNCTTTT 182
3236 rThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrV 3253
181 TGCTTTCAAACAGACACTGGCCTTTTGTGACGATCGCCTTCCTTGACT 132
3253 alThrProSerSerIle.....ProGlyThrThr 3262
131 TTCATCAGCGCGCATACCGCATATCTTCGCCAAGCAACGCGACTTC 82
3263 His....ThrAlaArgValLeuThrThrThrThrThrThrThrValAlaThrGI 3278
81 GGTAAATGACGGCGCGCTCAATAATGACT...TGCTCGGCTCTGCC 39
3278 ySerMetAlaThrProSerSerThrThrGlnThrSerGlyThrPro 3293
seq_name: pir2:T39903
seq_documentation_block:
serine-rich protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Nov-2000
C:Accession: T39903
R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Rieger, M.
submitted to the EMBL Data Library, November 1998
A:Reference number: Z21889
A:Accession: T39903
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-534 <L>N>
A:Cross-references: EMBL:AL033534; PIDN:CAA22127.1; GSPDB:GN00067; SPDB:SPBC215.13
A:Experimental source: strain 972h-; cosmid c215
C:Genetics:
A:Gene: SPDB:SPBC215.13
A:Map position: 2
C:Superfamily: pig submaxillary mucin
alignment_scores:
Quality: 121.00 Length: 420
Ratio: 0.582 Gaps: 15
Percent Similarity: 49.524 Percent Identity: 19.762
alignment_block:
US-09-303-518d-127/rev x T39903
Align seg 1/1 to: T39903 from: 1 to: 534
1280 GGGCAGACGAAGCTGCACAAAGAGCGGTCTTCTCGTCCATTCACAGCA 1231
```


99 CGAAGATATCGCGGTATCGCCCTNGATGAAAGTCAAGGAGCGCATG 148
51 userAsnHisAlaGlyAsnProAlaLysProValValSerProGlyAspG 68
149 CCGTCAAAAAAGAGCGAGTGTCTTGAAGACAAAAAGNATCCGGCGTG 198
68 lValLysThrGlyGlnValIleGlyGluProGlyGlyPheIleSerAla 84
199 GTGTTTACCGCGCGTTCAGGCAAAATCGCCCATTCATCGCGCGA 248
85 TyrLeuHisSerProValThrGlyArgValLeuGluIle..... 97
249 AAAGCGCTACTT.....CAGTCGTCTGTGA 274
98 .LysLysIleLeuHisProIleLeuGlyLysProIleGluAlaIleValI 114
275 TTGCGGTGAAGGCAACGACGAATC.....GAGTTC 306
114 leGluArgThrSerAspGluTrpValHisIleGluThrGlyAspPhe 130
307 GAACGCTACGCGCCGCAAGCGTTGCAAACTTA.....AGCGCGA 347
131 GluArgMetSerLysGluGluIleLeuGluIleLysLysAlaGlyI 147
348 NGAAATNNGNCAATCTCATCAATCCGTTTGTGGACTCGCGTGCCTA 397
147 e.....ValGlyLeuGlyGlyAlaMetPheP 156
398 NCCGTCGTTTCAAGAAATCCCTGCGTCGATCGCGAGCGCTCGCCATC 447
156 roThrHisValLysLeuSerProProGluLysLysValAspThrLeu 172
448 TTCGTCATGCGATGACACCAATCCGCTNGCGGACGACCCCTGTGTGT 497
173 ileValAsnGlyAlaGluCysGluProValLeuThrIleAspHisArgLe 189
498 GATCAAAAGAGCGCGANGATTTCAGACGANGTGTGTGTGTATTGAGCC 547
189 uMetLeuGluArgAlaGluAspIleLeuGlnGlyIleLeuIleMetMetL 206
548 GTTTCACCGAGCGTAAATCCATGCTGTAAAGGCA...GCTGGCGCAGAC 594
206 ysVal.....LeuGlyValGlnLysAlaValValGlyValGlu 218
595 GTGCGCTCTGAAATGCTCCCAACATCGAACACATGAATTCGCGCGCC 644
219 SerAsnLysMetAspAlaThrHisAsnLeuLysLysValPheLysGly.. 234
645 GCATCCGCGCGTGTGAGTGGCAGCAGCATTCATTTTCATTGAGCGCGTG 594
235 .TyrProValAspValAlaLeuLeuArgThrLysTyr.....ProGlnG 249
695 GTGCAACAAA...ACCGTTTGGACCATC..... 720
249 lAlaGluLysGlnLeuIleTyrAlaIleThrGlyArgMetValProArg 265
721AATTATCAAGATGT 734
266 GlyGlyLeuProMetAspValGlyValValValGlnAsnValGlyThrCy 282
735 AATTGCGATCGGACGTTTGTTCACACAGCGCTGTGAACACCGAGCGCG 784
282 sValAlaValLysGluAlaValValAspGlyLysProLeuValGluArgG 299
785 TGATTGCTTGGGTGTTCTCAAGTCAACAAACCCGCTCTTGGGTACC 834
299 lyMetThrValSerGlyAspAlaValLysAsnGlnLysAsnLeuIleVal 315
835 GTTTTGGGTGCGAAAGTATCGCAATTAAT.....GCGGCGCAATTGTT 878
316 ArgIleGlyThrProValLysAspValIleAspTyrCysGlyGlyIleAs 332

940 CACGATTATTGGGACGCTACCAACATCAGATTTCGTTATCGAAGAAG 989
331 331
990 CCGCAGCAAGAGAGCTGTTCGCTGGTGTGGCGCGCAGCGGACAAATACT 1039
332PrometMetGlyPheAlaLeuProAsp..... 340
1040 CCATCAGCGCTACGACCGCTCGGCCATTTCCTGAAACAAACACTTTCAG 1089
341LeuSerValProLeuIleLys 347
1090 TTCACGACCGCTCAACGCTGCGCGCGCCATGTGCGG..... 1131
348 ThrCysAsnCysLeuLeuAlaGlyAspAlaThrGluLeuProGluProVa 364
1132ATTGCTACTTACGAGCGGTAATCGCG 1159
364 lProAlaMetProCysIleArgCysGlyAspCysAlaGlnValCysProV 381
1160 TAGACATCTCGCTACCGCTGTTTGGCGGATTAATCTGCGGATACC 1209
381 alSerLeuLeuProGln...GlnLeuHisPhePheAlaLeuGlyAsp... 395
1210 GACAGCGCGCAAGCATGCGTGTGTTGGAATTGGACGAAGAAGACCTCGC 1259
396 GluHisGlnLeuLeuAlaHisAsnLeuPheAspCysIleGlyCysGI 412
1260 TTTGTCAGGTTCTGTCGCGCGCAATACGAATANGCGCGCTGTGC 1309
412 yAlaCysAlaTyrValCysProSerSerIle.....ProLeuValG 426
1310 GT 1311
426 In 426
seq_name: pir2:E72398
seq_documentation_block:
hypothetical protein TW0244 - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: E72398
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.
Nature 396, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316
A:Accession: E72398
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-451 <ARN>
A:Cross-references: GB:AE001708; GB:AE000512; NID:g4980740; PIDN:AAD35335.1; PID:g498074
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TW0244
alignment_scores:
Quality: 120.00 Length: 456
Ratio: 0.569 Gaps: 18
Percent Similarity: 46.272 Percent Identity: 20.175
alignment_block:
US-09-303-518d-127 x E72398 ..
Align seg 1/1 to: E72398 from: 1 to: 451
49 GAGCAAGTCATTATGACGGGCGCGCTCATTCACGAGTCGCTGTGCTGG 98
35 AspLysProIleGluArgAlaProLeuProGlnLysValPheValPheLe 51

675 AATGTGGTCCCACTCAAAACCGCGGATCCGGCGCCGCAATTCATGTC 626
227
625 TTTCGATGTTGGCAGCATTTTTCAGACGGCAGCTCTGGCCAGCTGCTTA 576
228 ..AlaThrThrAlaAlaThrThrAlaAlaThrThrAlaAla... 242
575 CACACATGGATTTTACGCTCGGTCAACGCGCTCAATACACAGCANACTCG 526
243ThrThrThrAlaAlaTh 248
525 TCTGAATCTGCGGCTCTTTTGATCAACACACAGGCTCTGCCGCA 476
248 rThrThrAlaAlaThrThrAlaAlaThrThrThrGlySerProThrS 265
475 GCGGATTTGCTTCATCGCATTCAGGAAGATGCGACAGCGCTGGCATCG 426
265 erGlySerThrSerThrThrGlyAlaSerThrSerThrProSerAlaSer 281
425 ACGGCA...GGGATTTGCTGACGGACGAGGATGCGACGCGCTGCCCAA 379
282 ThrAlaThrSerAlaThrProThrSerThrSerThrSerAlaAla.... 296
378 ACCGATTTGGATCAGATTGCGNCCNANTCTGCGCGCTTAAGTTTGCCA 329
297AlaThrThrSerThrProThrProThrProThrSerA 307
328 ACGCTTCGGGGCGGTAGCTTCGAACATCGATTTCGTGCTTCCTTCAACG 279
307 laAlaThrSerAla.....GluSerThrThrGluAlaProThrSer 320
278 GCAATCAGCAGCAGTCAAGTACGCGCTTTCGCGCGATGGATGGCGC 229
321 ThrProThrThrThrP...ThrThrThrProSerGluAlaThrThrAlaTh 336
228 GATTTTGGCTGAACGCGCGGTAAAC..... 201
336 rThrSerProGluSerThrThrValSerAlaSerThrThrSerAlaThrT 353
200ACACGCGCGGATNCTTTTGTCTTCAACAGACACTTGG 162
353 hrThrAlaPheThrThrGluSerHisThrSerProAspSerSerThrGly 369
161 CTTTTCGCGCATCGCTTCCTTCTGACTTCATCAGGGCGCATACC 112
370 SerThrSerThrAlaGluProSerSerThrPheThrLeu.....ThrPr 384
111 GGCATATTCTTCGCAACGACGACTTCGGTAATGACGGCGCGCTCAT 62
384 oSerThrAlaThrProSer.....ThrAspGlnPheThrGlySerSerA 399
61 AATGACT.....TCTCCGCTCTGCCCGCATGGC 30
399 laSerThrGluSerAspSerThrAspSerSerThrValProThrThrGly 415

seq_name: pir2:E64136

seq_documentation_block:

rnfc protein homolog - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 17-Mar-2000
C:Accession: E64136
R:Gocayne, J.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
; Fleischmann, R.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630
A:Accession: E64136
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA

A:Residues: 1-819 <TIGR>
A:Cross-references: GB:U32841; GB:L42023; NID:q1574529; PIDN:AAC2331.1; PID:q1574537
C:Superfamily: unassigned ferredoxin 2[4Fe-4S]-related proteins; ferredoxin 2[4Fe-4S]
F:371-435/Domain: ferredoxin 2[4Fe-4S] homology <FER>

alignment_scores:
Quality: 116.50 Length: 468
Ratio: 0.552 Gaps: 17
Percent Similarity: 45.085 Percent Identity: 19.017

alignment_block:

US-09-303-518d-127 x E64136 ..

Align seg 1/1 to: E64136 from: 1 to: 819

70 CCCGTCATTACGAAAGTCGGTTCCTGGCGAAGAAATATGCCGGTATGCG 119
38 ProLeuGlyThrAspPheTyrIleProLeuLysGlnHisLeuGlyThrTh 54
120 CCCCTNGATGAAAGTCAAGGAGCGATCGCGTCAAAAAAGGCCAAGTGC 169
54 rGlyAsnLeuLeuIleLysGluGlyAspTyrValLeuLysGlyGlnAla 71
170 TGTTCGAAACAAAAAGNATCCGGCGGTGGTTCACCGCCGCGTTCCTCA 219
71 euThrLysGlyAspGlyLeuArgMetLeuProValHisAlaProThrSer 87
220 GCGAAATTCGCGCGCATCCATCCGCGCGGAAAGCGGTACTTCATCGCT 269
88 GlyThrIle.....LysSerIleLysProTyrVa 97
270 CGTGATTGCGGTGAAGCAACGACGCAAAATCGAG..... 303
97 lAlaThrHisProSerGlyLeuAspGluProThrIleHisLeuGlnAla 114
304TTCGACGCTAGCGCCGCAAGCGGTGGCAAC 336
114 spGlyLeuAspGlnTrpIleGluArgAsnProIleAspAspSerThr 130
337 TTAACGCGCGGANGAANTNNGCAATCTGATCCAAATCGCGTTCGTGGAC 386
131 LeuSerSerGluGlnLeuIleHisIleTyrGlnAlaGlyIle...Al 146
387 TCGCTGCTGCTANCGTCCGCTTC.....ACGAAATTCCTCGCGTCG 427
146 aGlyLeuGlyGlyAlaValPheProThrAlaAlaLysIleGlnSerAla 163
428 ATCCGAGCGGTTCGCCATCTTCGTCATCGCATGGACACCAATCCGCTN 477
163 luGlnLysValLysLeuLeuIleAsnGlyAlaGluCysGluProTyr 179
478 GCGGACAGCCCTGTGTTGTGATCAAGAGCGCGANGANGATTTCAGACG 527
180 IleThrCysAspAspArgLeuMetArgGluArgAlaAspGluIleIle 196
528 ANGTTGCTGGTATTGAGCGGTTTGACGAGCGGTAAATCCATGCTGTA 577
196 sGlyIleArgIleLeuArgTyrIleLeuHisProGlu...LysValVal 212
578 AGCAGCTGCGCGACGTCGCGTCT..... 603
212 leaIleGluAspAsnLysProGluAlaIleSerAlaIleArgAsnAla 228
604 ...GAAATGCTCCCAACATCAAAACACATGAATTCGGCGCCCGCATCC 650
229 LeuGlnGlyAlaAsnAspIleSerIleArgValIleProThrLysTyrPr 245
651 GCGCGGTTGAGTGGCAGCAGCATTCATTTCATTGAG.....C 688
245 oSerGlyAlaThrLysGlnLeuIleTyrLeuLeuThrGlyIleGluValP 262
689 CGGTGCGTCAACAAACCGGTTTGGACCATCAATTATCAAGATGTAATT 738

R; Ricke, D.O.; Bruce, D.; Mundt, M.; Doggett, N.; Munk, C.; Saunders, E.; Robinson, D.;
 re, J.; White, S.; Teng, S.; Tatum, O.; Campbell, C.; Fawcett, J.; Deaven, L.
 submitted to the EMBL Data Library, March 1998
 A: Description: Sequencing of human chromosome 16p13.3.
 A: Reference number: Z14564
 A: Accession: T02345
 A: Status: preliminary; translated from GB/EMBL/DBJ

A:Reference number: I52257; MUID:92068178
A:Accession: I52257
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-631 <RES>
A:Cross-references: GB:M77226; NID:g199835; PIDN:AAA39754.1; PID:g199837
A:Accession: I65210
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-631 <RE2>
A:Cross-references: GB:M84683; NID:g199842; PIDN:AAA39756.1; PID:g199843
C:Genetics:
A:Gene: Muc1
A:Introns: 20/1; 454/3; 472/2; 517/1; 557/3; 607/3

alignment_scores:

Quality:	109.00	Length:	470
Ratio:	0.534	Gaps:	19
Percent Similarity:	43.404	Percent Identity:	22.766

alignment_block:

US-09-303-518D-127/rev x I52257

Align seg 1/1 to: I52257 from: 1 to: 631

```
1343 CAGCCTCTCTCTCAANGTTTCCAGCACCTTA.....CG 1309
|||||  ::  |||:::|||||
35 GluAspThrSerSerSerLeuAlaSerThrThrProValHisSerSe 51
1308 CAACAGCGGGCCTATTCTGTTTCCGCGGAGAGCTGCACAAAG 1259
|||||  |||  ::  |||:::|||||
51 rAsnSerAspProAlaThrArgProGlyAspSer..... 63
1258 CGAGGTCTTCTTCGTCCTCAATCCAGCAACCAATGCTTCCGCGCTGCG 1209
64 .....:::..Thr 64
1208 GTATCGCGCAGATTAATTCGCGCAAAAGCAGGTTAGCAGGATCTCTAG 1159
|||||  ::  |||:::|||||
65 SerSerProValGlnSerSerThrSerProAlaThrArgAlaProG1 81
1158 CGGCATTACGCGCTCGTAACTTGAAGAGTGTCTTTTTCAGGAAATGGCCG 1059
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
81 uAspSerThrSerThrAlaValLeuSerGlyThr.....SerSerProA 96
1108 CGTTCAGCGCTGCTGTAAGTGTGTAAGAGTGTCTTTTTCAGGAAATGGCCG 1059
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
96 laThrThrAlaProValAsnSerAlaSer..... 105
1058 AGGTCGTACGCGGTGATGAGTATTGTCGCGCTGCGCGCAACCCAGCC 1009
|||||  ::  |||:::|||||
106 .....:::..SerProValAlaHisGlyAspThrSerSerPr 116
1008 GAACAGCTCTTGTGTCGCGCTCTCTCCATAACGGAATCTGATGTGGT 959
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
116 caLathrSerLeuSerLysAspSer..... 124
958 AGCGTCCAAAATAATCGTCGCGCTGTGTGTAATCGCGCGCTTCAATACC 909
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
125 .....:::..AsnSerSerProValValHisSerGlyThrSerSer 136
908 GAACCGGAAATACAGCGGTGCTGCGTCAACCAATTCGCGCGCAGTAAT 859
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
137 AlaProAlaThrThrAlaProValAspSerThrSerSerProValValH 153
858 TTGCGATACTTTCACACCAAAAGGTACCAAGAGCGGTGTTGTTGA 809
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
153 sGlyGlyThrSerSerProAlaThr.....SerProGlyAspSerT 168
808 CTTGAGAACCAACCAACCAATCAGCGCTGCTGTTTCAGACGCGCTGTT 759
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
168 hrSerSerProAspHisSerThrSerSer.....Pro 179
```

```
758 GCAACAAACGTCCTGATGGCAATACATCTTATATAATTTGATGTCCTCAAC 709
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
180 AlaThrArgAlaProGluAspSerThrSer.....Th 190
708 GGTGTTGTTTTCACGACCGCTCAATGAATGAATGATGCTGCTGCCACTCA 659
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
190 rAlaValLeuSerGlyThrSerSer..... 198
658 AACCGCGGATGCGCGCGCGCAATTCATGTTTTCGATGTTTGGCA... 612
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
199 ..ProAlaThrThrAlaProValAspSerThrSerProValAlaHis 214
611 .....GCATTTTCAGAC..... 600
215 AspAspThrSerSerProAlaThrSerLeuSerGluAspSerAlaSerSe 231
599 .....GGCAGTCTGCGCGCAGCTGCTTACACACATGGA 566
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
231 rProValAlaHisGlyGlyThrSerSerProAlaThr..... 243
565 TTTTACGCTCGTCAACGCTCAATACACAGCANACNTCTCTGAAATCN 516
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
244 .....SerProLeuArgAspSerThrSerProValHisSerSer 257
515 TCGNCGGCTTCTTTCATCACACACAGGCTGCTGCCGCGNAGC..... 474
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
258 AlaSerIleGlnAsnIleLysThrSerAspLeuAlaSerThrProAs 274
473 ....GGATTGCTGCTCATGACGAGATGCGCAAGATGCGGCTGCGCAT 428
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
274 pHisAsnGlyThrSerValThrThrSerSerAlaLeuGlySerAla 291
427 CGACGCGAGGATTTGCTGAACGAGCGGNTACGACGCGAGTCCACAAA 378
291 hrSer..... 292
377 CCGATTGTCATGATTCGNNCNCNNTTCGCGCTTAAGTTTGCCTAA 328
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
293 ProAsp.....HisSerGlyThrSerThrThrThrAs 303
327 CGCTTCGCGCGCTGAGCTTTCGAACTCG.....ATTTCGCTGCTGCT 285
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
303 nSerSerGluSerValLeuAlaThrThrProValTyrSerSerMetProp 320
284 ..TCAACGCGCAATCACACGCTGAGTACGCGCTTTTCGCGCGCATGG 237
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
320 heSerThrThrLysValThrSerGlySerAlaIleProAspHisAsn 336
236 ATGCGCGGATTTTGCCT..... 219
337 GlySerSerValLeuProThrSerSerValLeuGlySerAlaThrSerLe 353
218 .....GAAACNGCGCGTAAACACGCGCGCTTCTTTTGTCTT 176
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
353 uValTyrAsnThrSerAlaIleAlaThrThrPro.....ValS 366
175 CAACACGCTTTCGCTTTTTCACGCGATCGCTTCTTCTGATTTTCATC 126
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
366 erAsnGlyThrGlnProSerValProSerGlnTyrProValSerProThr 382
125 NAGCGCGCATACCGGCTATTTCTGCGCAACGACGCGCTTCGGTAAT 76
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
383 MetAlaThrThrSerSerHisSerThrIleAlaSerSerTyrTyrSe 399
75 GACGGCGCGC 66
399 rThrValPro 402
seq_name: pir2-C39135
seq_documentation_block:
hypothetical protein 3 (gyrB region) - Haloferax sp.
```

C; Species: Haloferax sp.
C; Date: 11-Feb-1993 #sequence_revision 11-Feb-1993 #text_change 18-Nov-1994

C; Accession: C39135
R; Holmes, M.L.; Dyall-Smith, M.L.

J. Bacteriol. 173, 642-648, 1991

A; Title: Mutations in DNA gyrase result in novobiocin resistance in halophilic archaeobac

A; Reference number: A39135; MUID:91100352

A; Accession: C39135

A; Status: preliminary; translation not shown

A; Molecule type: DNA

A; Residues: 1-437 <HOL>

A; Cross-references: GB:M38373

alignment_scores:
Quality: 108.00 Length: 480
Ratio: 0.554 Gaps: 27
Percent Similarity: 40.625 Percent Identity: 23.958

alignment_block:

US-09-303-518d-127 x C39135 ..

Align seg 1/1 to: C39135 from: 1 to: 437

```

12 CAATAAGGTCTAAACCTGCCATCGCGGAGACCGGAGCAAGTCAATTT 61
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
48 GlnAspAlaArgLysProProAlaArgAlaArgValProGlnLeu.... 62
62 ATGACGGCCGCTCATACCAAGTGGTGGTGGCGAAGATATGCC 111
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
63 ....ArgGlyArgAspPheAlaLeuArgArgAlaAspArgVal...G 77
112 GGTATCGCCCTGTGATGAAAGTCAAGGAGCGATGCCGTCACAAAGG 161
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
77 LuHisValProLeu.....ArgGlyArgHisProArgValArgArg 90
162 ...CCAAGTGCT.....GTTTGAAG 178
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
91 ValProGlnArgAspGlnAspGlyAlaProArgArgHisLeuLeuAr 107
179 ACAAAGNATCCGG...CGTGGTGTATACCGC..... 209
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
107 gArgArgValGlyGlyHisArgGlyArgAsnArgHisAlaGlyAspArg 124
210 .....GCCNGTTTCAGGCACAAATCGCGCCATCCATCG 242
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
124 rgAlaProGlyValAspSerArgLeuArgGlnGlnHisGlnHisProArg 140
243 CGCGCAAGACCGCTACTTCAGTCCGTGCTGATTCGCTTGAAGCAACG 292
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
141 GlyArgHisAlaSerAspArgValGlnAspGlyAla.....HisProArg 155
293 ACGAATCGAGTTCGAACGCTACGCGCCCGGAGCGTGGCAAACTTAAGC 342
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
155 gArgGlnArgLeuArgGlu...GlnProArgHisAlaGlyArgProArg 171
343 GCGGANGAANTNNGNCAATGTATCCAAATCCGCTTTGTGACTGCGCT 392
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
171 rgArg..... 172
393 GCGTANCCGCTTCAGCAAAATCCCTGCCGTCGATGCCGAGCC..... 437
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
173 .....GlnProArgArgArgGlyArgSerArgGln 182
438 .....GTTCCGCACTTCGTCAATGCGATGGACACCAATCCGCTNCGC 480
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
182 yThrHisArgArgHisLeuArgGlnAlaPro..... 192
481 GCAGACCCCTGT.....GGTGTGATCAAGAGACGCGNGANGATTTCAG 524
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
193 ..ArgProAlaValArgGlyProAspGluAspGlnAlaArgGlnPheArg 208
525 ACGANGTNGCTGCTATTCAGCGGTTGACCGCGTTGACCGCGCTAAATCCATGCT 574

```

```

208 ..... 208
575 GTAAGGCAGCTGGCGCAGACAGCTGCCGTCTGAAATGTCTGCCAACATCGAA 624
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
209 ....GlyProArgHisArgArgGlu..... 215
625 ACACATGAATTCGGCGCGCCGATCCGCCGCGG.....TTT 659
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
216 .....ArgHisProProThrAlaArgAspValLeuArgGlyGlu 228
660 GAGTGGCAGCAGCATTCATTTCATTCAGCGCGTGGTGGTGGTGGTTCCAAGT 709
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
229 ProGlyHisGlyAspGlyHisLeuGluGlyArgArgGlyArgProAr 245
710 TTGGACCATCAATTATCAAGATGTAATTCATTCAGCGCGTGGTGGTGGTTC 759
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
245 gProGlnGlyArgGluAlaGlyArgGlyAlaHis..... 256
760 ACAGCGCTCTGAACACCGCGCGTGGTGGTGGTGGTGGTGGTTCCAAGT 809
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
257 .....ProProGlnValArgAlaArgGlyLeuAlaAlaGlyGluAla 271
810 CAA.....CAAAACACCGCTCTT.....GCGTACCGTTT 838
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
272 ArgGlyLeuProGluProArgProLeuGlyValArgThrValHisArgG 288
839 TGGTGGCAAGATTCGCAAAATTCGCGCGCGAATGTTGGTGGTGGTGGTGG 888
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
288 yGly.....ArgLeuArgGlyArgValGlnAlaGlyProArgProG 303
889 AA...CCGCGTGATTC.....CGTGGTGGT 911
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
303 InValProGlyAspPheAlaProGlnGlyGluAspSerGluArgArgG 319
912 ATTTGAACCGCGATTCACACAGCGCGCGCAGCATTTTGGTGGTGGTGGT 961
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
320 ThrProArgProHisSerArgLysArgArgAspThrGlyAlaHisH 336
962 ACAATCAGATTCGTTATCGAAGAGCGCGCAGCAAGAGCTGTTCGCGC 1011
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
336 sArgHisTrpArgArgArgArgArgValArgHisArgGlyGluGlyAla 353
1012 TGGGTTTCGCGCGCGCGGACAAATCTCCATCAGCGCTACGACCGCTCG 1061
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
353 eu.....ProAla 355
1062 CCATTTCTGAAACAAACTCTTCAAGTTCACGACAGCGCTACCGTG 1111
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
356 AlaHisProAsp.....AspArgArgArgArgArg 365
1112 GCGACCGCGCGATGTCGCGATTTGGTACTTACGAGCGGTAATGCGCGCTA 1161
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
365 gArgArgAlaHis.....ProAspAlaAla 374
1162 GACATCTGCTACCTCGCTTTTTCGCGGATTTAATCGTGGCGATACCGA 1211
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
374 lArgAlaSerValProAlaHisAlaProAlaHisArgGlyArgLeuArg 390
1212 CAGCGCGCAAGCATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1261
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
391 ValArgGlySerThrAla.....AlaValProArgProLeuProArg.. 404
1262 TGTGACGCTGCTGCGCGCGGCAAAATACGAATANGGCC 1301
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
405 ..GlnHisLeuArgArgAspGlyArgGlyArgAlaGlyPro 417

```

seq_name: pir2:T45134

seq_documentation_block:
hypothetical protein [imported] - Microbacterium ammoniaphilum (fragment)
C; Species: Microbacterium ammoniaphilum
C; Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000

C:Accession: T45134
 R:Striebel, H.M.; Seiber, S.; Jarsch, M.; Kessler, C.
 Gene 172, 41-46, 1996
 A:Title: Cloning and characterization of the Mami restriction-modification system from M
 A:Reference number: 222923; MUID: 96257250
 A:Accession: T45134
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-529 <STR>
 A:Cross-references: EMBL:X79027; NID:9984667; PIDN:CAA55649.1; PID:g1679831
 A:Experimental source: ATCC 15354

alignment_scores:
 Quality: 108.00 Length: 473
 Ratio: 0.548 Gaps: 24
 Percent Similarity: 41.649 Percent Identity: 24.101

alignment_block:

US-09-303-518d-127 x T45134 ..

Align seg 1/1 to: T45134 from: 1 to: 529

```

27 COTGCCCATCGCGGCACACCGGACCAAGTCAATTAAGCGGCGCCGPCA 76
|||||:|||||:
128 ProAlaArgArgGlyValLeuArgArgHisArgHisArgAlaArgG1 144
77 TTACCGAAGTCGCGTGTGCGGCAAGATATGCGCGTATCGCGCCCTNG 126
|||||:|||||:
144 yGluArgGlyArgGlyProArgGlnGlnValProArgGlnHisPro... 159
127 ATGAAGTCAGGAAGGCGATGCGGTCAAAAGGCCAAGTCGTGTTGA 176
|||||:|||||:
160 .....ArgGlyArgArgAspArgAlaGlyArgProGlyLeuHisArg 173
177 AGACAAAAGNATCC .....GGCGTGGTGTATACCGCGCNG 214
|||||:|||||:
174 ArgArgArgAlaArgArgGlnGlnArgProGlnValArgHisG1 190
215 TTTCAGCGCAAAATCGCGCCATCATCGCGGCAAGCGGTACTTCAG 264
|||||:|||||:
190 yAspGlnHisArgAlaAspProArgArgProArgAspProArgAlag 207
265 TCGGTGCTGATGCGGTGAAGCAACGACGAATCGAGTCGACGCTA 314
|||||:|||||:
207 lyHisHisProLeuArgGlnGlyGlnGlnGluAlaArgProLeu 223
315 CCGCGCCGCAAGC .....GTTGGCAAACTTAAGCGCGCANG 349
|||||:|||||:
224 ArgAlaArgGlyGlnGlnGlyProGlyGlyProAlaGlyArgHisPr 240
350 AANTNNGNCAATCTGATCCAAATCCGTTTGTGGACTGCGCTGCTANC 399
|||:|||||:
240 oAlaLeuGlyAspGluasp .....Argp 248
400 CGTCCGTTACGAAATCCCTGCGTGCATGCGGAGCCGTTCCGCACTT 449
|||||:|||||:
248 roArgAlaArgGlnGlyAlaArgAlaAlaHisArgGluAlaLeuHis 264
450 CGTCAATGCGATGACACCAATCCGTCNGCGGACGACCCGTGTGTTGT 497
|||||:|||||:
265 ArgLeu .....GlnArgArgGlyArgAlaAspGlyCysAr 277
498 ....GATCAAGAGCGCGANGATTTTCAGACGANGTNTGCTGGTATTG 543
|||:|||||:
277 gAlaGlnGlyArgAlaArgAlaArgAlaArgAlaGlyGlyGlyLeu 293
544 AGCGGTTTGACGAGCGCTAAATCCATGTGTGTAAGCAGCTGCGCAGA 593
|||||:|||||:
294 ProArgArgGluaspArg ..... 299
594 CGTGCCGCTGTAATGCTGCCAACATCGAAACACATGAATTCGGCGGC 643
|||||:|||||:

```

```

300 .....ValArgAlaHisArg .....P 305
644 GGCATCCGCGCGGTTTGAAGTGGCAGCAGCATTCATTTTCATTTGAGCGGTC 693
|||||:|||||:
305 roArgProArgGlyArgGlyGlyAlaAlaArgLeuAspArgAlaGlyG1 321
694 GGTCAAAACAAACCGTTTGGACCAATATCAAGATGTAATTCGCAT 743
|||||:|||||:
321 yValGlyAlaArgProAlaArgProArgArgLeuArgHisProArgProA 338
|||||:|||||:
744 CGGACGTTTGTTCGACAGCGCTCTGAACACCGGCGGTGATTGCTT 793
|||||:|||||:
338 laAspLeuProHisGlyGlyPro ..... 345
794 TGGGTGTTTCTCAAGTCAACAAACCGCCTCTTGGCTACCGTTTGGGT 843
|||||:|||||:
346 .....GlnGlyValAlaArgLeuAspHisProGlnGlyL 357
844 GCGAAAGTATCGCAAAATTAAGTGGCGGGAATTTGGTTCAGCGCAGCAACCG 893
|||||:|||||:
357 euGlu .....GlyAlaAlaGlySerArgArgHisPro 367
894 CGTGATTTCGCGTTTCGCTATTGAA .....CGGCGGATTTACACAAAGCG 937
|||||:|||||:
368 HisArgLeuArgAlaArgLeuHisGlnGlyArgGlyAspLeuLeuArgAr 384
938 CGCACGATTAATTTGGAGCGCTACCAATCAGATTTCCCTTTATCGAAGAA 987
|||||:|||||:
384 g .....ProArgArgAspArgLeuGlyArgArg 394
988 GGCGCGCAGCAAGAGCTGTTCGCTGGTTCGCGCGCAGCGCGCAGCAATA 1037
|||||:|||||:
394 lyProArgGlnGlyGln .....GlyAlaHisGlyGlyGlnGly 406
1038 CTCATCAGCGCTACGACCTCGGCCATTTCTGAAAAACAACTCTTCA 1087
|||||:|||||:
407 LeuArgHisAla ..... 410
1088 AGTTCAGCAGACCGCTCAACGGTGGCGACCG .....C 1119
|||||:|||||:
411 .....GlyArgArgArgGlyValProSerGlnLeuAlaValLeuA 425
1120 GCATGCT .....GCCGATTGGTACTTACGA 1145
|||||:|||||:
425 rgGlnGlyValCysGlnAlaSerAlaThrLeuAlaLeuTrpMetThrSer 441
1146 GCGCGTAATGCGCGTAGACATCTGCTACCTGCTTTTTCGCGGATTTAA 1195
|||||:|||||:
442 GlyArg...LeuLeuArgGluLeuAlaGlyLeuLeuCysArgAspServ 457
1196 TCGTCGCGCATACCGAC .....AGC 1215
|||||:|||||:
457 alGluSerHisLeuAspValThrTrpArgAlaAspAlaSerIleSer 473
1216 GCGCAAGCATTTGGTTCCTTGAATTTGGAGAAAGAACACCTCGCTTGTG 1265
|||||:|||||:
474 .ArgArgThrTrpThrValMetAsn...ThrArgLeuArgGlnAlaGlnL 489
1266 CAGCTTCGTCCTCCGCGCAATACGAATANGCCCGCTGTTCGTAAGG 1315
|||||:|||||:
489 ysSerLeuAlaAlaArgProThrThrSerSerIleSerArgSer 505
1316 TGCTGGAAA 1324
|||||:|||||:
506 CysTrpGln 508
seq_name: pir2:T45462
seq_documentation_block:
membrane glycoprotein [imported] - equine herpesvirus 1
C:Species: equine herpesvirus 1
C:Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 09-Jun-2000
C:Accession: T45462

```


[illegible]

527 pValCysLeuValGIuAAsnLysProGIuAlaThrLysCysValAlaCysG 54

663 ACTCAAAACGGCGGATCGGGCGCGCAATTCATGTGTTCCATGTTGG 614

544 luThrSerLysProGIuThrLysAlaGIuLeuLysGlyPheGlyThr.s 560

613 CAGCATTTTCAGACGGCAGCTCTCGCCAGCTGCCTTACACACATGGAT 564

560 erThrPheSerSerGIyThrAlaAlaProThr..... 570

563 TTACGCTCGGTCAAACGGCTCAATACACAGCANCTCGTCTGAATCN.. 516

571 PheLysPheGlyValGIuInSerSerAspSerThrAlaGIuLeuLysSerGI 587

515 ...TCGNCGGCTTCTTTGTATGATCAACAACAGGCTGTCCGCNAGCGGAT 470

587 yAlaSerThrSerGIyPheAlaLysSerIleGIyAspPheLysPheGIyL 604

469 TGGTGTCCATCGCATTCAGGAAGATGCGCAACGGCTCGCGATCG..... 426

604 euValSerAlaSerThrThrThrGIuGIuThrLysLysSerPheThr 620

425ACGGCAGGATTTTGTCTGAACGG 403

621 PheGIySerSerThrThrAsnGIuValSerAlaGIyPheLysPhe...GI 636

402 ACGGNTAGCAGCGGATGCCACAACCGGATTCGATCGATTCGCNNCNA 353

636 yIleAlaGIySerAlaGIuInThrLysProAspThrLeuSer...GlnSerT 652

352 NTTCTNTCCGCTTAAGTTTGCACAACGCTTCGGCGCGTTCGCTCGAAC 303

652 hrThrSerGIyPheThrPheGIySerVal.....SerAsn 663

302 TCGATTTTCGTCGTTTCCTTCACGGCAATCAGCCGACTGAAGTACGCG 253

664 ThrValSerLeuAlaPro.....AlaAlaThrSerSerSerThrGI 678

252 CTTTTCGCGCGATGGATGGCGGATTTTGCCTGAAACNGCGCGGTAA 203

678 yLeuGIu.....ValAlaAlaIleAlaAspSerAsnLeuAlaT 692

202 ACAC..... 198

692 hrThrAlaAlaLeuLysSerAlaGIuLysLysAlaGIuAlaProThr 708

197 ...ACGCGCGGATNCTTTTGTCTCAACACAGACATTCGCCCTTTTTCAC 151

709 IleThrProPheSerPheGIyLysThrAspGIuAsn.....Ly 721

150 GGCATTCGCTTCTTGACTTTCATCNAAGGGCGCATACCGCATATTCCT 101

721 sGIuThrAlaSerThrSerPheValPheGIyLys.....L 733

100 CGCCAAGCAACGCCACTTCGGTAAATGACGGCCGTCATAAATGACTTCG 51

733 ysAspGIuLysThrAspSerAlaProThrGIySerSer.....PheAla 747

50 TCCGGTGTGCGCGGATGGGC...AGSTTTAGACCTTTTTCG 12

```
alignment_scores:
  Quality: 107.50
  Ratio: 0.846
Length: 254
Gaps: 11
```

Percent Similarity: 50.000 Percent Identity: 24.803

alignment_block:

US-09-303-518D-127/rev x A41258

Align seg 1/1 to: A41258 from: 1 to: 725

```
618 GTTGCAGCATTTTCAGACGGCAGCTCTGCGCAGCTGCCTTACACACAT 569
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
124 ValThrSerLysPheThrSerTyrIleCys...HisThr 139
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
568 GGATTTTACGCTCGCTCAACAGCGCTCAATACC.....AGCANA 531
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
139 hrAlaIleSerSerLeuSerGluValGlyThrThrValValSerSer 155
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
530 CNTCTCTGAAATNTCCGCGCTTCTTTGATC..... 498
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
156 SerAlaIleGluProSerSerAlaSerIleIleSerProValThrSerTh 172
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
497 .....ACAA 494
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
172 rLeuSerSerThrThrSerSerAsnProThrThrThrSerLeuSerSerT 189
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
493 CCACAGGCTCTGCCNAGCGGATGTTGTCATCGCATTCACGAAGATG 444
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
189 hrSerThrSerProSerSerThrSerThrSerProSerSerThrSerThr 205
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
443 GCGAACGGCTCGGCATCGACGGCAGGATTTTGCTGAACGGACGGTACG 394
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
206 SerSerSerSerThrSerThrSerSerSerSerThrSerThrSerSerSe 222
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
393 CAGCCAGCTCCACAAACCGGATTTGGATCAGA.....TTGC 359
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
222 rSerThrSerThrSerProSerSerThrSerThrSerSerSerLeuThrS 239
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
358 NNCNNANTCNTCCGCTTAAGTTTGCAACGCTTCGGCGCGCTAGCGT 309
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
239 erThrSerSerSerSerThrSerThrSerGlnSerSerThrSerThrSer 255
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
308 TCGAACTCGATTTCGTGCTTCCTTCAACGGCAATCACACCGACTGAAG 259
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
256 SerSerSerSerThrSerSerProSerSerThrSerThrSerSerSerSe 272
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
258 TACGGCTTTTCGGCGGATGGATGGCGGATTTTCCTGAAACNGGCG 209
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
272 rThrSerThrProSer.....SerLysSerThrSerA 284
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
208 CGGTAACACACACCGCGGATNCTTTTGTCTTCAACACAGCATTTGCCT 159
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
284 laSerSerThrSerThrSerSerTyr.....SerThrSerThrSerPro 298
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
158 TTTTTCAGCGCA.....TCGCTTCCTTCGAC 133
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
299 SerLeuThrSerSerSerProThrLeuAlaSerThrSerProSerSerTh 315
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
132 TTTTCATNAGGGGGGCGATA..... 114
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
315 rSerIleSerSerThrPheThrAspSerThrSerSerLeuGlySerSerI 332
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
113 .....CCGCA 108
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
332 leAlaSerSerThrSerSerValSerLeuTyrSerProSerThrProVal 348
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
107 TATTCCTTCGCAACGACGCGACTTCGGTAATGACGGGCCCGCATAAAT 58
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
349 TyrSerValProSerThrSerSerAsnValAlaThr...ProSer...Me 363
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
57 GACTTGTCTCC 48
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
363 tThrSerSer 366
```

seq_name: pir2:B46629

seq_documentation_block:

mucin 6, gastric (3-repeat clone) - human (fragment)

N:Alternate names: MUC6

C:Species: Homo sapiens (man)

C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999

C:Accession: B46629

R:Toribara, N.W.; Robertson, A.M.; Ho, S.B.; Kuo, W.L.; Gum, E.; Hicks, J.W.; Gum Jr., J. Biol. Chem. 268, 5879-5885, 1993

A:Title: Human gastric mucin. Identification of a unique species by expression cloning

A:Reference number: A46629; MUID:93194895

A:Accession: B46629

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-505 <TOR>

A:Cross-references: GB:L07518; MID:g292045; PIDN:AAB61945.1; PID:g292046

A:Experimental source: stomach

A:Note: sequence extracted from NCBI backbone (NCBIN:128397, NCBIP:128399)

C:Genetics:

A:Gene: GDB:MUC6

A:Cross-references: GDB:134734; OMIM:158374

A:Map position: l1p15.5-l1p15.5

C:Keywords: glycoprotein

alignment_scores:

Quality: 107.00 Length: 310

Ratio: 0.690 Gaps: 14

Percent Similarity: 50.000 Percent Identity: 22.258

alignment_block:

US-09-303-518D-127/rev x B46629

Align seg 1/1 to: B46629 from: 1 to: 505

926 ATCCGCGCTTCAATACCGAATCACGCGGTGTCTGCGTCAAC 877

||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

87 ValAlaProThrSerThr...ThrThrIleThrProAsnProThrSerTh 102

||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

876 CAATTCCGCGCAGTAATTTGCGATACTTTCGCACCCCAAAACGGTACGCA 827

||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

102 rArgThrArgThrProValAlaHisThrAsnSerAlaThrSerSerArgp 119

||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

826 AGAGCGGTGTTTGTGACTTGAGAACCCACCAAGCAATCACGCGCTCG 777

||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

119 roProProPheThrThrHisSerProProThrGlySerSer..... 133

||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

776 GTGTTTCAGAGCGCTTTCGCAACAAACGTCGATGGCAATTACATCTTG 727

||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

134ProPheSerSerThrGlyProMetThrAlaThrSer... 145

||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

726 ATAATTGATGTCACAAACGGTTTGTTCACCGCGCTCAATGAAT 677

||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

146PheLysThrThrThrTyrProThrProSerLeu.... 157

||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

676 GAATGTGCTGCCACTCAAAACCGCGGATCGCGGCCGCGCAATTTCATGT 627

||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

158ProGlnThrThrPro 162

||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

626 GTTTCGATGTTGCAGCATTTTCAGACGCGCATCTGCGCCAGCTGCTT 577

||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

163 LeuThrHisValProPheSerThrSerLeuValThrProIleThr... 178

||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

576 ACACACATGGATTTACGCTCGGTCAACGCGCTCAATACCAACAGCANA.... 531

||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

179 .HisThrValIleThrProThrHisProGlnMetSerThrSerAlaTyrI 195

||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

530CNTGCTGCTGAATTCN 516

||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

195 leHisSerThrProThrGlyThrIleAlaSerProThrThrValLysAla 211

||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

515 TCGNCGGCTTCTTGTATCAACACACAGGCTGCTCCGCGNACGGATTCGT 466

||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

212 ThrArgSerThrThrAlaProLeuMetThrAlaThrThrSerArgIl 228

||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

seq_name: pirl:PWPFBL

seq_documentation_block:

H+-transporting two-sector ATPase (EC 3.6.3.14) beta chain - brown alga (*Pyraiaella littoralis*)
 C:Species: chloroplast Pyraiaella littoralis
 C:Date: 31-Mar-1992 #sequence-revision 31-Mar-1992 #text_change 14-Dec-2001
 C:Accession: S20848; S17102
 R:Jouanin, S.; Kerbourc'h, C.; Kloareg, B.; Loiseaux-de Goer, S.
 Plant Mol. Biol. 18, 819-822, 1992
 A:Title: Nucleotide sequences of the atpB and the atpF genes of the brown alga *Pyraiaella littoralis*
 A:Reference number: S20848; NUID:92216062

A:Accession: S20848

A:Molecule type: DNA

A:Residues: 1-481 <OU>

A:Cross-references: EMBL:X60329; NID:g14180; PIDN:CAA42899.1; PID:g14181

C:Genetics:

A:Gene: atpB

A:Genome: Chloroplast

C:Superfamily: H+-transporting ATP synthase alpha chain; H+-transporting ATP synthase alpha chain; ATP biosynthesis; chloroplast; hydrolase; membrane-associated complex; nucleotide-binding motif A (P-loop)
 F:161-168/Region: nucleotide-binding motif A (P-loop)
 F:187-363/Domain: H+-transporting ATP synthase alpha chain homology <ATP>

alignment_scores:

Quality:	106.50	Length:	510
Ratio:	0.493	Gaps:	25
Percent Similarity:	42.353	Percent Identity:	22.157

alignment_block:

US-09-303-518D-127 x PWPFBL

Align seg 1/1 to: PWPFBL from: 1 to: 481

```

13  AAAAAAGTCTAAACCTGCCATCGCGGACACCGGAGCAAGTCATTTA 62
   |||||
9  GlulysGlyleasn.....AsnGlyTyrIleThrGlnValIle.. 21

63  TGACGGCCCGTCATACCGAAGTC.....GCGTTCCTGGCGAAG 103
   |||||
22  ...GlyProValIleAspAlaValPheSerSerGlyIleLeuProLysI 37

104 AATATCGCGGTATCGCCCTGATGAACTCAAGGAGCGATGCGCGTC 153
   |||||
37  IeYrAsnAlaLeuGlu.....ValGlnSerLysGluGlyProIleIle 51

154 AAAAAAGCCCAAGTGTGTGTAAGACAAAAAGNATCCGGCGCTGTGTT 203
   |||||
52  CysGluValGlnGlnLeuLeuGlyAspAsnArgValArgAlaIleAla 68

204 TACGGCCCGTTCAGGCAAAATCGCGCCATCCATCGCGGCAAAAGC 253
   |||||
68  tSerAla.....ThrAspGlyLeuGlnArgGlyValThrV 80

254 GCGTACTTCAGTCGTCGTGATTCGCGTTCAA..... 285
   |||||
80  allleaspthrGlnAlaProIleAlaValProValGlyLysAlaThrLeu 96

285 ..... 285

97  GlyArgIlePheAsnValLeuGlyGlnProValAspAsnLeuSerAsp 113

286 .....GGCAACGCAAGATCGAGTTCGACGCTACGCGCCCGGAGCGTTGG 331
   |||||
113 rValGlyGluAspThrLeuProIleHisArgSerAlaPro...AlaPhe 129

332 CAAACTTAAGCGCGGANGAANTNNGCAATCTGATCAATCCGTTTG 381
   |||||
129 hrAspLeu.....GluThrLysProAlaIlePheGluThrGlyIle 142

382 TGGACTCGCGTTCGCTGATTCGCTGACGCAAAATCCCTGCGCTCGATGC 431
   |||||
143 LysValValAspLeuLeuAlaProTyrArgArg.....Gln 154

```

```

432 CGAGCCGTTCCGCGCATCTTCGTCATCGGATGGACACCAATCCGCTNGCGG 481
   |||||
154 YglyIleGlyLeuPheGlyGlyAla.....GlyV 165

482 CACACCCCTGTGGTGTGATCAAGAAGCCGCGGANGATTCAGACGANGT 531
   |||||
165 alGlyIsthrrValLeuIleMetGluLeuAsnAsnIleAlaLys... 180

532 NTGCTGGTATTAGCGCGTTTGGACGCGTAAATCCATCTGTGTAAAGC 581
   |||||
181 .....AlaHisGlyGlyValSerValPheGlyG 190

582 AGCTGGCGGACGAGTCCCTCTGAAATGCTGCCACATCGCAACACATG 631
   |||||
190 YValGlyGluArgThrArgGluGlyAsnAspLeuTyrMetGluMetLysG 207

632 AATTCGGCGCGCGCATCGCGCGGTTTGGTGGCAGCACATTCATTTC 681
   |||||
207 luSerGlyValIleAsnGluThrAsnLeuLeuGluSerLysValAlaLeu 223

682 AYT.....GAGCGGTCGGTGCACAAACAAACCGTTGG... 714
   |||||
224 ValTyrGlyGlnMetAsnGluProGlyAlaArgMetArgValGlyLe 240

715 .....ACCATCAATTCATCAAGATG 733
   |||||
240 uThrAlaLeuThrMetAlaGluTyrPheArgAspIleAsnLysGlnAspV 257

734 TAATTGCC...ATCGGACGTTTGTTCACACGCGCTGTGAACACCGAG 780
   |||||
257 alLeuLeuPheIleAspAsnIlePheArgPheValGlnAlaGlySerGlu 273

781 CCGTTCATTGCTTTGGGT.....GGTCTCAAGTCACAAACACCGCCT 824
   |||||
274 ValSerAlaLeuLeuGlyArgMetProSerAlaValGlyTyrGlnProTh 290

825 CTTGCGTACCGCTTTGGGTGCGAAAGTATCGCAAAATTCATCGCGCGA 874
   |||||
290 rLeuGlyThrGluMetGlyAlaLeuGlnGluArgIleThr..... 303

875 TGGTTGACGACAGCAACCGCGTATTCGCGTTCGTTGATTTGAACGCGCG 924
   |||||
304 .....SerThrThrGlnGlySer 309

925 ATTACACAGGCGCGACGATTTATTTGGACGCTACCAATCATCATTC 974
   |||||
310 IleThrSerIleGlnAlaValTyrVal..... 318

975 CGTTATCGAAGAGCGCGACGCAAAAGAGTGTTCGCTGGTGGTTCGCGCG 1024
   |||||
319 .....ProAlaAspAspLeu.....ThrAspProA 327

1025 ACCCGGCAAAATCTCATCGCGCTACGACCCCTGGCGCATTCCTGAAA 1074
   |||||
327 laPro.....AlaThrThrPheAlaHis..... 334

1075 AACAACTCTCAAGTTCACGACACCGCTCAACGCTGCGGACCGCGCAT 1124
   |||||
335 .....LeuAspAlaThrThrValLeuSer.....ArgGlyLe 345

1125 GGTGCGCGATTGCTACTTACGAGCGGTAATCGCGCTAGACATCTGCGCTA 1174
   |||||
345 uAlaAlaLysGlyIleTyrProAlaValAspProLeuAspSerThrSert 362

1175 CCCTGCTTTGCGGATTAATCGTCGCGAT..... 1206
   |||||
362 hr...MetLeuGlnProLeuIleValGlyAspGluHisTyrLysThrAla 377

1207 .....ACCGACACGCG 1217
   |||||
378 GlnLeuValLysGluThrLeuGlnArgTyrLysGluLeuGlnAspIleI 394

```


720 yserThrGluserThrValThrGlySerSerValSerThrValSerGly 736
